```
sites: std
length: 4102 (circular)

            pleI
            hinfI
          xhoI  salI
          paeR7I taqI              aluI
     ecoRI taqI hincII/hindII  tru9I                  eco57I                                                    mnlI
     apoI avaI  accI   aciI    mseI                  bslI                                          fokI     ddeI
   1 GAATTCTCGA GTCGACGGCG GGCGTTAAAG CTCTCGTGGC ATTATCCTTC AGTGGGGCTA TTGGACTGAC TTTTCTTATG CTGGGATGTG CCTTAGAGGA
     CTTAAGAGCT CAGCTGCCGC CCGCAATTTC GAGAGCACCG TAATAGGAAG TCACCCCGAT AACCTGACTG AAAAGAATAC GACCCTACAC GGAATCTCCT

                                                                                                          tru9I
            rsaI                                                                                          mseI
            csp6I    eco57I                  apoI              maeIII     apoI                          mseI
 101 TTATGGGTGT ACTTCTCTGA AGTAAGATGA TTTGTCAAAA ATTCTGTGTG GTTTTGTTAC ATTGGGAATT TATTTATGTG ATAACTGCGT TTAACTTGTC
     AATACCCACA TGAAGAGACT TCATTCTACT AAACAGTTTT TAAGACACAC CAAAACAATG TAACCCTTAA ATAAATACAC TATTGACGCA AATTGAACAG
   1                         M I   C Q K   F C V   V L L H   W E F   I Y V   I T A F   N L S

                                          nlaIII
                                         sphI
                                         nspI                                                             pleI
                  styI       tru9I       nspHI                                                            hinfI                apoI
       xcmI       bsaJI      mseI          apoI                                                                                 apoI
 201 ATATCCAATT ACTCCTTGGA GATTTAAGTT GTCTTGCATG CCACCAAATT CAACCTATGA CTACTTCCTT TTGCCTGCTG GACTCTCAAA GAATACTTCA
     TATAGGTTAA TGAGGAACCT CTAAATTCAA CAGAACGTAC GGTGGTTTAA GTTGGATACT GATGAAGGAA AACGGACGAC CTGAGAGTTT CTTATGAAGT
  26 Y P I   T P W R   F K L   S C M   P P N S   T Y D   Y F L   L P A G   L S K   N T S

      taqI
      sfuI
      bstBI                 pvuII
      bsiCI                 nspBII                  tru9I        rsaI
      asuII            bsmAI aluI        ddeI      mseI         csp6I
 301 AATTCGAATG GACATTATGA GACAGCTGTT GAACCTAAGT TTAATTCAAG TGGTACTCAC TTTTCTAACT TATCCAAAAC AACTTTCCAC TGTTGCTTTC
     TTAAGCTTAC CTGTAATACT CTGTCGACAA CTTGGATTCA AATTAAGTTC ACCATGAGTG AAAAGATTGA ATAGGTTTTG TTGAAAGGTG ACAACGAAAG
  59 N S N G   H Y E   T A V   E P K F   N S S   G T H   F S N L   S K T   T F H   C C F R

                                                                                  apoI                    sfaNI
 401 GGAGTGAGCA AGATAGAAAC TGCTCCTTAT GTGCAGACAA CATTGAAGGA AAGACATTTG TTTCAACAGT AAATTCTTTA GTTTTTCAAC AAATAGATGC
     CCTCACTCGT TCTATCTTTG ACGAGGAATA CACGTCTGTT GTAACTTCCT TTCTGTAAAC AAAGTTGTCA TTTAAGAAAT CAAAAAGTTG TTTATCTACG
  93   S E Q   D R N   C S L C   A D N   I E G   K T F V   S T V   N S L   V F Q Q   I D A
```

FIG. 1A

```
                                                                                  xmnI
                                                                                  tfiI
                                                                                  hinfI                     sau96I
                                                                  pleI      tru9I                           avaII
                                             tru9I                hinfI     mseI asp700                     asuI
       bsrI                            bsmAI mseI
501  AAACTGGAAC ATACAGTGCT GGCTAAAAGG AGACTTAAAA TTATTCATCT GTTATGTGGA GTCATTATTT AAGAATCTAT TCAGGAATTA TAACTATAAG
     TTTGACCTTG TATGTCACGA CCGATTTTCC TCTGAATTTT AATAAGTAGA CAATACACCT CAGTAATAAA TTCTTAGATA AGTCCTTAAT ATTGATATTC
126   N  W  N   I  Q  C  W   L  K  G   D  L  K   L  F  I  C   Y  V  E   S  L  F   K  N  L  F   R  N  Y   N  Y  K

                                                                                                          nlaIII
                                tfiI                                                                      rcaI
                                hinfI  mnlI      bslI                                          munI       bspHI
                      eco57I  mboII hphI      nlaIV
601  GTCCATCTTT TATATGTTCT GCCTGAAGTG TTAGAAGATT CACCTCTGGT TCCCCAAAAA GGCAGTTTTC AGATGGTTCA CTGCAATTGC AGTGTTCATG
     CAGGTAGAAA ATATACAAGA CGGACTTCAC AATCTTCTAA GTGGAGACCA AGGGGTTTTT CCGTCAAAAG TCTACCAAGT GACGTTAACG TCACAAGTAC
159  V  H  L  L   Y  V  L   P  E  V   L  E  D  S   P  L  V   P  Q  K   G  S  F  Q   M  V  H   C  N  C   S  V  H  E

                                                                                                          hphI
                                                                                                          maeIII
                                                                                                     bsrI    mnlI
701  AATGTTGTGA ATGTCTTGTG CCTGTGCCAA CAGCCAAACT CAACGACACT CTCCTTATGT GTTTGAAAAT CACATCTGGT GGAGTAATTT TCCAGTCACC
     TTACAACACT TACAGAACAC GGACACGGTT GTCGGTTTGA GTTGCTGTGA GAGGAATACA CAAACTTTTA GTGTAGACCA CCTCATTAAA AGGTCAGTGG
193   C  C  E   C  L  V   P  V  P  T   A  K  L   N  D  T   L  L  M  C   L  K  I   T  S  G   G  V  I  F   Q  S  P

                                              bslI
                                              sau3AI
                                              mboI/ndeII[dam-]
                                              dpnI[dam+]                                                  xcmI
                                              dpnII[dam-]                                   tru9I         sau96I
                                              alwI[dam-]                                    mseI          avaII
                   bstXI        hphI                       ndeI                             ahaIII/draI   asuI
801  TCTAATGTCA GTTCAGCCCA TAAATATGGT GAAGCCTGAT CCACCATTAG GTTTGCATAT GGAAATCACA GATGATGGTA ATTTAAAGAT TTCTTGGTCC
     AGATTACAGT CAAGTCGGGT ATTTATACCA CTTCGGACTA GGTGGTAATC CAAACGTATA CCTTTAGTGT CTACTACCAT TAAATTTCTA AAGAACCAGG
226   L  M  S   V  Q  P  I   N  M  V   K  P  D   P  P  L  G   L  H  M   E  I  T   D  D  G  N   L  K  I   S  W  S
```

FIG. 1B

```
                   rsaI
                    csp6I
                  nlaIV
                  kpnI
                  hgiCI
                  banI
                  asp718                             ecoRI                      drdI        ddeI
      bslI        acc65I                 sspI       apoI                  aluI        bsmAI aluI fokI
 901 AGCCCACCAT TGGTACCATT TCCACTTCAA TATCAAGTGA AATATTCAGA GAATTCTACA ACAGTTATCA GAGAAGCTGA CAAGATTGTC TCAGCTACAT
     TCGGGTGGTA ACCATGGTAA AGGTGAAGTT ATAGTTCACT TTATAAGTCT CTTAAGATGT TGTCAATAGT CTCTTCGACT GTTCTAACAG AGTCGATGTA
 259 S  P  P  L   V  P  F   P  L  Q   Y  Q  V  K   Y  S  E   N  S  T   T  V  I  R   E  A  D   K  I  V   S  A  T  S
                                                              ^begin12u

                                                                                 tfiI
                                                                               scrFI
                                                                               mvaI
                                                                               ecoRII
                                                                               dsaV
                               bpuAI                                      bstXI
                            scrFI                                              bstNI
                            mvaI bbsI                                    apyI[dcm+]
                            ecoRII                                   sau96I              rsaI
                            dsaV                                       haeIII/palI           bsrI   csp6I
          accI              bstNI                     mnlI          bsrI  asuI     hinfI     maeIII  scaI
        rmaI      bst1107I  bsaJI mboII              hphI      bsmAI  fokI  bsaJI    gsuI/bpmI gsuI/bpmI  mnlI
        maeI      accI      apyI[dcm+]      mnlI
1001 CCCTGCTAGT AGACAGTATA CTTCCTGGGT CTTCGTATGA GGTTCAGGTG AGGGGCAAGA GACTGGATGG CCCAGGAATC TGGAGTGACT GGAGTACTCC
     GGGACGATCA TCTGTCATAT GAAGGACCCA GAAGCATACT CCAAGTCCAC TCCCCGTTCT CTGACCTACC GGGTCCTTAG ACCTCACTGA CCTCATGAGG
 293  L  L  V   D  S  I   L  P  G  S   S  Y  E   V  Q  V   R  G  K  R   L  D  G   P  G  I   W  S  D  W   S  T  P

                                                 apoI                                             sfaNI
1101 TCGTGTCTTT ACCACACAAG ATGTCATATA CTTTCCACCT AAAATTCTGA CAAGTGTTGG GTCTAATGTT TCTTTTCACT GCATCTATAA GAAGGAAAAC
     AGCACAGAAA TGGTGTGTTC TACAGTATAT GAAAGGTGGA TTTTAAGACT GTTCACAACC CAGATTACAA AGAAAAGTGA CGTAGATATT CTTCCTTTTG
 326  R  V  F   T  T  Q  D   V  I  Y   F  P  P   K  I  L  T   S  V  G   S  N  V   S  F  H  C   I  Y  K   K  E  N
```

FIG. 1C

```
                                                                                              nlaIII
                                                                                             sau3AI
                                                                                             mboI/ndeII[dam-]
                                                                                             dpnI[dam+]
                                                                                             dpnII[dam-]
                                               apoI    ddeI                                  bclI[dam-]         maeIII
                   mnlI                     fokI       aluI      apoI  mnlI     bsrI
     1201 AAGATTGTTC CCTCAAAAGA GATTGTTTGG TGGATGAATT TAGCTGAGAA AATTCCTCAA AGCCAGTATG ATGTTGTGAG TGATCATGTT AGCAAAGTTA
          TTCTAACAAG GGAGTTTTCT CTAACAAACC ACCTACTTAA ATCGACTCTT TTAAGGAGTT TCGGTCATAC TACAACACTC ACTAGTACAA TCGTTTCAAT
      359 K  I  V  P   S  K  E   I  V  W   W  M  N  L   A  E  K   I  P  Q   S  Q  Y  D   V  V  S   D  H  V   S  K  V  T

                               taqI
                              xhoI
                              paeR7I
                              avaI                                                          bsmI
                              mnlI mnlI                 rsaI   fnu4HI                  nlaIII
                                              sfaNI     csp6I  bbvI
     1301 CTTTTTTCAA TCTGAATGAA ACCAAACCTC GAGGAAAGTT TACCTATGAT GCAGTGTACT GCTGCAATGA ACATGAATGC CATCATCGCT ATGCTGAATT
          GAAAAAAGTT AGACTTACTT TGGTTTGGAG CTCCTTTCAA ATGGATACTA CGTCACATGA CGACGTTACT TGTACTTACG GTAGTAGCGA TACGACTTAA
      393  F  F  N   L  N  E   T  K  P  R   G  K  F   T  Y  D   A  V  Y  C   C  N  E   H  E  C   H  H  R  Y   A  E  L

                                                      tru9I                                 rsaI
                                                      mseI                                  csp6I
                                                   rsaI                             bsrI              maeIII
                                                   csp6I                      hincII/hindII    bsrI           aciI
                                      nlaIII
     1401 ATATGTGATT GATGTCAATA TCAATATCTC ATGTGAAACT GATGGGTACT TAACTAAAAT GACTTGCAGA TGGTCAACCA GTACAATCCA GTCACTTGCG
          TATACACTAA CTACAGTTAT AGTTATAGAG TACACTTTGA CTACCCATGA ATTGATTTTA CTGAACGTCT ACCAGTTGGT CATGTTAGGT CAGTGAACGC
      426  Y  V  I   D  V  N  I   N  I  S   C  E  T   D  G  Y  L   T  K  M   T  C  R   W  S  T  S   T  I  Q   S  L  A

                                                                                         hgiJII
                                                                                         bsp1286
                                                                                         bmyI
                                           fnu4HI                                        banII
                                           bbvI                            fokI         ddeI
                    munI mnlI
     1501 GAAAGCACTT TGCAATTGAG GTATCATAGG AGCAGCCTTT ACTGTTCTGA TATTCCATCT ATTCATCCCA TATCTGAGCC CAAAGATTGC TATTTGCAGA
          CTTTCGTGAA ACGTTAACTC CATAGTATCC TCGTCGGAAA TGACAAGACT ATAAGGTAGA TAAGTAGGGT ATAGACTCGG GTTTCTAACG ATAAACGTCT
      459 E  S  T  L   Q  L  R   Y  H  R   S  S  L  Y   C  S  D   I  P  S   I  H  P  I   S  E  P   K  D  C   Y  L  Q  S
```

FIG. 1D

```
                                                                                sau3AI
                                                                                mboI/ndeII[dam-]
                                                                                dpnI[dam+]
                                                                                dpnII[dam-]
                                                                               alwI[dam-]
                     ppu10I                                                mamI[dam-]              rmaI        pleI
                     nsiI/avaIII                                 draIII    bsaBI[dam-]             maeI        hinfI
                   bsmI                  mboII
1601 GTGATGGTTT TTATGAATGC ATTTTCCAGC CAATCTTCCT ATTATCTGGC TACACAATGT GGATTAGGAT CAATCACTCT CTAGGTTCAC TTGACTCTCC
     CACTACCAAA AATACTTACG TAAAAGGTCG GTTAGAAGGA TAATAGACCG ATGTGTTACA CCTAATCCTA GTTAGTGAGA GATCCAAGTG AACTGAGAGG
493   D  G  F   Y  E  C   I  F  Q  P   I  F  L   L  S  G   Y  T  M  W   I  R  I   N  H  S   L  G  S  L   D  S  P
                           ^begin13-2

        nlaIII
       nspI
       nspHI                                          fokI
       aflIII            tfiI                       mnlI  bsrI
                         hinfI     hphI
1701 ACCAACATGT GTCCTTCCTG ATTCTGTGGT GAAGCCACTG CCTCCATCCA GTGTGAAAGC AGAAATTACT ATAAACATTG GATTATTGAA AATATCTTGG
     TGGTTGTACA CAGGAAGGAC TAAGACACCA CTTCGGTGAC GGAGGTAGGT CACACTTTCG TCTTTAATGA TATTTGTAAC CTAATAACTT TTATAGAACC
526  P  T  C   V  L  P  D   S  V  V   K  P  L   P  P  S  S   V  K  A   E  I  T   I  N  I  G   L  L  K   I  S  W

                                     tfiI
                                     hinfI           tru9I         rsaI
                                  xcmI               mseI          csp6I    mboII       mnlI       sfaNI
       bsrI
1801 GAAAAGCCAG TCTTTCCAGA GAATAACCTT CAATTCCAGA TTCGCTATGG TTTAAGTGGA AAAGAAGTAC AATGGAAGAT GTATGAGGTT TATGATGCAA
     CTTTTCGGTC AGAAAGGTCT CTTATTGGAA GTTAAGGTCT AAGCGATACC AAATTCACCT TTTCTTCATG TTACCTTCTA CATACTCCAA ATACTACGTT
559 E  K  P  V   E  P  E   N  N  L   Q  F  Q  I   R  Y  G   L  S  G   K  E  V  Q   W  K  M   Y  E  V   Y  D  A  K

                                                            hinPI           rmaI
                                                            hhaI/cfoI mnlI  maeI      bsrI
                  bsmAI bsrI
1901 AATCAAAATC TGTCAGTCTC CCAGTTCCAG ACTTGTGTGC AGTCTATGCT GTTCAGGTGC GCTGTAAGAG GCTAGATGGA CTGGGATATT GGAGTAATTG
     TTAGTTTTAG ACAGTCAGAG GGTCAAGGTC TGAACACACG TCAGATACGA CAAGTCCACG CGACATTCTC CGATCTACCT GACCCTATAA CCTCATTAAC
593   S  K  S   V  S  L   P  V  P  D   L  C  A   V  Y  A   V  Q  V  R   C  K  R   L  D  G   L  G  Y  W   S  N  W

                                                  sau96I
                                                  avaII
                                                  asuI
                                                 ppuMI                         tru9I
                                          ecoNI  eco0109I/draII                mseI
                           nlaIII         bslI  mnlI      apoI                aseI/asnI/vspI
2001 GAGCAATCCA GCCTACACAG TTGTCATGGA TATAAAAGTT CCTATGAGAG GACCTGAATT TTGGAGAATA ATTAATGGAG ATACTATGAA AAAGGAGAAA
     CTCGTTAGGT CGGATGTGTC AACAGTACCT ATATTTTCAA GGATACTCTC CTGGACTTAA AACCTCTTAT TAATTACCTC TATGATACTT TTTCCTCTTT
626   S  N  P   A  Y  T  V   V  M  D   I  K  V   P  M  R  G   P  E  F   W  R  I   I  N  G  D   T  M  K   K  E  K
```

FIG. 1E

```
                                        pleI
                                        hinfI                                                                  nlaIII
        maeIII
2101 AATGTCACTT TACTTTGGAA GCCCCTGATG AAAAATGACT CATTGTGCAG TGTTCAGAGA TATGTGATAA ACCATCATAC TTCCTGCAAT GGAACATGGT
     TTACAGTGAA ATGAAACCTT CGGGGACTAC TTTTTACTGA GTAACACGTC ACAAGTCTCT ATACACTATT TGGTAGTATG AAGGACGTTA CCTTGTACCA
 659 N  V  T  L   L  W  K   P  L  M   K  N  D  S   L  C  S   V  Q  R   Y  V  I  N   H  H  T   S  C  N   G  T  W  S

                                                                  haeIII/palI
                                                                mscI/balI
                                                                haeI
                                                                eaeI
                                                    maeIII      cfrI          munI
      mboII                   apoI
2201 CAGAAGATGT GGGAAATCAC ACGAAATTCA CTTTCCTGTG GACAGAGCAA GCACATACTG TTACGGTTCT GGCCATCAAT TCAATTGGTG CTTCTGTTGC
     GTCTTCTACA CCCTTTAGTG TGCTTTAAGT GAAAGGACAC CTGTCTCGTT CGTGTATGAC AATGCCAAGA CCGGTAGTTA AGTTAACCAC GAAGACAACG
  693   E  D  V   G  N  H   T  K  F  T   F  L  W   T  E  Q   A  H  T  V   T  V  L   A  I  N   S  I  G  A   S  V  A

                                                                                                        scrFI
                                                                                                        mvaI
                                                                                                        ecoRII
                                                                                                        dsaV
                                                                                                        bstNI
                                                                                                        apyI[dcm+]
                        bslI      haeIII/palI                                ddeI          tru9I
                 tru9I tru9I       haeI                                     draIII        mseI
     apoI mseI   mseI              nlaIII                                  maeIII      ahaIII/draI
2301 AAATTTTAAT TTAACCTTTT CATGGCCTAT GAGCAAAGTA AATATCGTGC AGTCACTCAG TGCTTATCCT TTAAACAGCA GTTGTGTGAT TGTTTCCTGG
     TTTAAAATTA AATTGGAAAA GTACCGGATA CTCGTTTCAT TTATAGCACG TCAGTGAGTC ACGAATAGGA AATTTGTCGT CAACACACTA ACAAAGGACC
 726  N  F  N   L  T  F  S   W  P  M   S  K  V   N  I  V  Q   S  L  S   A  Y  P   L  N  S  S   C  V  I   V  S  W

                                                                                                  xmnI
                                                                                                  tfiI
                                                                                                  hinfI mboII
                                                                                                  asp700
                                                                tru9I
                                                                mseI  mboII hphI                ddeI    earI/ksp632I
          draIII
          hphI bsrI          aluI
2401 ATACTATCAC CCAGTGATTA CAAGCTAATG TATTTTATTA TTGAGTGGAA AAATCTTAAT GAAGATGGTG AAATAAAATG GCTTAGAATC TCTTCATCTG
     TATGATAGTG GGTCACTAAT GTTCGATTAC ATAAAATAAT AACTCACCTT TTTAGAATTA CTTCTACCAC TTTATTTTAC CGAATCTTAG AGAAGTAGAC
 759  I  L  S  P   S  D  Y   K  L  M   Y  F  I  I   E  W  K   N  L  N   E  D  G  E   I  K  W   L  R  I   S  S  S  V
```

FIG. 1F

```
                          sau3AI
                          mboI/ndeII[dam-]
                          dpnI[dam+]
                          dpnII[dam-]                    bsrI                                                                  tru9I
                                                                                                                               mseI
     tru9I                bclI[dam-]                     rsaI                                                                  aseI/asnI/vspI
     mseI                 nlaIII                         csp6I                        sspI
2501 TTAAGAAGTA TTATATCCAT GATCATTTTA TCCCCATTGA GAAGTACCAG TTCAGTCTTT ACCCAATATT TATGGAAGGA GTGGGAAAAC CAAAGATAAT
     AATTCTTCAT AATATAGGTA CTAGTAAAAT AGGGGTAACT CTTCATGGTC AAGTCAGAAA TGGGTTATAA ATACCTTCCT CACCCTTTTG GTTTCTATTA
 793   K  K  Y   Y  I  H   D  H  F  I   P  I  E   K  Y  Q   F  S  L  Y   P  I  F   M  E  G   V  G  K  P   K  I  I

                                                                                                 mboII
                                                                                                 earI/ksp632I
                                                    bspMI                                        mnlI
                                                    sfaNI                        bsrI
2601 TAATAGTTTC ACTCAAGATG ATATTGAAAA ACACCAGAGT GATGCAGGTT TATATGTAAT TGTGCCAGTA ATTATTTCCT CTTCCATCTT ATTGCTTGGA
     ATTATCAAAG TGAGTTCTAC TATAACTTTT TGTGGTCTCA CTACGTCCAA ATATACATTA ACACGGTCAT TAATAAAGGA GAAGGTAGAA TAACGAACCT
 826   N  S  F   T  Q  D  D   I  E  K   H  Q  S   D  A  G  L   Y  V  I   V  P  V   I  I  S  S   S  I  L   L  L  G

                                                                                                 bsp1286
                                                                                                 bmyI
                                                                                           scrFI
                                                                                           mvaI
                                                                                           ecoRII
                                                                                           dsaV
                                                                                           bstNI
          tru9I                                     xmnI                                   bsaJI                 tru9I
          mseI                                      mboII                                  apyI[dcm+]            mseI
          aseI/asnI/vspI                 aluI       asp700
2701 ACATTATTAA TATCACACCA AAGAATGAAA AAGCTATTTT GGGAAGATGT TCCGAACCCC AAGAATTGTT CCTGGGCACA AGGACTTAAT TTTCAGAAGC
     TGTAATAATT ATAGTGTGGT TTCTTACTTT TTCGATAAAA CCCTTCTACA AGGCTTGGGG TTCTTAACAA GGACCCGTGT TCCTGAATTA AAAGTCTTCG
 859 T  L  L  I   S  H  Q   R  M  K   K  L  F  W   E  D  V   P  N  P   K  N  C  S   W  A  Q   G  L  N   F  Q  K  P

                                                    nlaIII
                                                    nspI       earI/ksp632I
                                                    nspHI  sau96I
                                                    aflIII avaII
                                                    eam1105I   mnlI                                                        ecoRV
          maeII                                                                                                           mboII
          psp1406I  sfaNI                                sfaNI  maeIII     asuI mboII      nlaIV
2801 CAGAAACGTT TGAGCATCTT TTTATCAAGC ATACAGCATC AGTGACATGT GGTCCTCTTC TTTTGGAGCC TGAAACAATT TCAGAAGATA TCAGTGTTGA
     GTCTTTGCAA ACTCGTAGAA AAATAGTTCG TATGTCGTAG TCACTGTACA CCAGGAGAAG AAAACCTCGG ACTTTGTTAA AGTCTTCTAT AGTCACAACT
 893   E  T  F   E  H  L   F  I  K  H   T  A  S   V  T  C   G  P  L  L   L  E  P   E  T  I   S  E  D  I   S  V  D
```

FIG. 1G

```
                                                                    sau3AI
                                                                    mboI/ndeII[dam-]
                                                                    dpnI[dam+]
                                                                    dpnII[dam-]
                                                bsmAI               bstYI/xhoII                                bsrI
        nlaIII                   sfaNI          bsaI                bglII                                   maeIII
2901 TACATCATGG AAAAATAAAG ATGAGATGAT GCCAACAACT GTGGTCTCTC TACTTTCAAC AACAGATCTT GAAAAGGGTT CTGTTTGTAT TAGTGACCAG
     ATGTAGTACC TTTTTATTTC TACTCTACTA CGGTTGTTGA CACCAGAGAG ATGAAAGTTG TTGTCTAGAA CTTTTCCCAA GACAAACATA ATCACTGGTC
926   T  S  W   K  N  K  D   E  M  M   P  T  T   V  V  S  L   L  S  T   T  D  L   E  K  G  S   V  C  I   S  D  Q

                                                                                                       sau3AI
                                                                                                       mboI/ndeII[dam-]
             tru9I                                                                                     dpnI[dam+]
             mseI                 ddeI maeIII                                                tru9I     dpnII[dam-]
           hpaI        ddeI   ddeI   rsaI mnlI                            bsmAI               mseI      bclI[dam-]
           hincII/hindII mnlI  mnlI  csp6I  bstEII     mnlI
3001 TTCAACAGTG TTAACTTCTC TGAGGCTGAG GGTACTGAGG TAACCTATGA GGACGAAAGC CAGAGACAAC CCTTTGTTAA ATACGCCACG CTGATCAGCA
     AAGTTGTCAC AATTGAAGAG ACTCCGACTC CCATGACTCC ATTGGATACT CCTGCTTTCG GTCTCTGTTG GGAAACAATT TATGCGGTGC GACTAGTCGT
959 F  N  S  V   N  F  S   E  A  E   G  T  E  V   T  Y  E   D  E  S   Q  R  Q  P   F  V  K   Y  A  T   L  I  S  N

                                                         draIII
                      hphI                               hphI              rmaI                         tfiI
                   bsrI  mboII                           maeIII            maeI      apoI               hinfI
3101 ACTCTAAACC AAGTGAAACT GGTGAAGAAC AAGGGCTTAT AAATAGTTCA GTCACCAAGT GCTTCTCTAG CAAAAATTCT CCGTTGAAGG ATTCTTTCTC
     TGAGATTTGG TTCACTTTGA CCACTTCTTG TTCCCGAATA TTTATCAAGT CAGTGGTTCA CGAAGAGATC GTTTTTAAGA GGCAACTTCC TAAGAAAGAG
993   S  K  P   S  E  T   G  E  E  Q   G  L  I   N  S  S   V  T  K  C   F  S  S   K  N  S   P  L  K  D   S  F  S

                       scrFI
                       mvaI
                       ecoRII
                       dsaV
                       bstNI                  sfaNI
                       apyI[dcm+]             sau3AI fokI
                     sau96I                   mboI/ndeII[dam-]
                     haeIII/palI           mamI[dam-]
                     asuI                     dpnI[dam+]
        nlaIII                                dpnII[dam-]
     aluI          mnlI bsaJI                 bsaBI[dam-]        hphI      mnlI      ddeI       fokI
3201 TAATAGCTCA TGGGAGATAG AGGCCCAGGC ATTTTTTATA TTATCAGATC AGCATCCCAA CATAATTTCA CCACACCTCA CATTCTCAGA AGGATTGGAT
     ATTATCGAGT ACCCTCTATC TCCGGGTCCG TAAAAAATAT AATAGTCTAG TCGTAGGGTT GTATTAAAGT GGTGTGGAGT GTAAGAGTCT TCCTAACCTA
1026 N  S  S   W  E  I  E   A  Q  A   F  F  I   L  S  D  Q   H  P  N   I  I  S   P  H  L  T   F  S  E   G  L  D
```

FIG. 1H

```
                                                                                                                    mnlI
                                                                                                                  hphI
                                                                                                                maeIII
                                                                                                               bstEII
                                                                mboII
                                                mnlI  apoI    eco57I
3301 GAACTTTTGA AATTGGAGGG AAATTTCCCT GAAGAAAATA ATGATAAAAA GTCTATCTAT TATTTAGGGG TCACCTCAAT CAAAAAGAGA GAGAGTGGTG
     CTTGAAAACT TTAACCTCCC TTTAAAGGGA CTTCTTTTAT TACTATTTTT CAGATAGATA ATAAATCCCC AGTGGAGTTA GTTTTTCTCT CTCTCACCAC
1059 E  L  L  K   L  E  G   N  F  P   E  E  N  N   D  K  K   S  I  Y   Y  L  G  V   T  S  I   K  K  R   E  S  G  V

                                                                                                                scrFI
                                                                                                                mvaI
                                                                                                                ecoRII
                                                                                                                dsaV
                                                                                                                bstNI
                                                                                                                apyI[dcm+]
                                                                bsp1286                                       gsuI/bpmI
                                drdI                            bmyI
3401 TGCTTTTGAC TGACAAGTCA AGGGTATCGT GCCCATTCCC AGCCCCCTGT TTATTCACGG ACATCAGAGT TCTCCAGGAC AGTTGCTCAC ACTTTGTAGA
     ACGAAAACTG ACTGTTCAGT TCCCATAGCA CGGGTAAGGG TCGGGGGACA AATAAGTGCC TGTAGTCTCA AGAGGTCCTG TCAACGAGTG TGAAACATCT
1093   L  L  T   D  K  S   R  V  S  C   P  F  P   A  P  C   L  F  T  D   I  R  V   L  Q  D   S  C  S  H   F  V  E

                                                                                                                nlaIII
                                                                                                              sau3AI
                                                                                                                mboI/ndeII[dam-]
                                                                                                      pleI      dpnI[dam+]
                                                                                                      hinfI     dpnII[dam-]
                                                                mboII                nlaIII
                                                        rmaI    bpuAI               nspI
                                        ddeI            maeI    bbsI       sfaNI    nspHI mnlI                  ddeI
3501 AAATAATATC AACTTAGGAA CTTCTAGTAA GAAGACTTTT GCATCTTACA TGCCTCAATT CCAAACTTGT TCTACTCAGA CTCATAAGAT CATGGAAAAC
     TTTATTATAG TTGAATCCTT GAAGATCATT CTTCTGAAAA CGTAGAATGT ACGGAGTTAA GGTTTGAACA AGATGAGTCT GAGTATTCTA GTACCTTTTG
1126  N  N  I   N  L  G  T   S  S  K   K  T  F   A  S  Y  M   P  Q  F   Q  T  C   S  T  Q  T   H  K  I   M  E  N

                                                                mboII
                                maeIII                        eco57I     eco57I
3601 AAGATGTGTG ACCTAACTGT GTAATTTCAC TGAAGAAACC TTCAGATTTG TGTTATAATG GGTAATATAA AGTGTAATAG ATTATAGTTG TGGGTGGGAG
     TTCTACACAC TGGATTGACA CATTAAAGTG ACTTCTTTGG AAGTCTAAAC ACAATATTAC CCATTATATT TCACATTATC TAATATCAAC ACCCACCCTC
1159 K  M  C  D   L  T  V

                                                                                                                xmnI
                                pleI                                                                          asp700
                                hinfI apoI                                                ddeI maeIII   apoI
3701 AGAGAAAAGA AACCAGAGTC AAATTTGAAA ATAATTGTTC CAAATGAATG TTGTCTGTTT GTTCTCTCTT AGTAACATAG ACAAAAAATT TGAGAAAGCC
     TCTCTTTTCT TTGGTCTCAG TTTAAACTTT TATTAACAAG GTTTACTTAC AACAGACAAA CAAGAGAGAA TCATTGTATC TGTTTTTTAA ACTCTTTCGG
```

FIG. 1I

```
                                                                                                  sau96I
                                                                                                  nlaIV
                                                                                                  avaII
                                                mboII                          rmaI               asuI                    rmaI
                                                earI/ksp632I                   maeI             ppuMI                     maeI
                                    accI      sapI                        aluI                  eco0109I/draII    aluI
3801 TTCATAAGCC TACCAATGTA GACACGCTCT TCTATTTTAT TCCCAAGCTC TAGTGGGAAG GTCCCTTGTT TCCAGCTAGA AATAAGCCCA ACAGACACCA
     AAGTATTCGG ATGGTTACAT CTGTGCGAGA AGATAAAATA AGGGTTCGAG ATCACCCTTC CAGGGAACAA AGGTCGATCT TTATTCGGGT TGTCTGTGGT

                                                                                                                        nspI
                                                                                                                        nspHI
                                                                                                                  tru9I nlaIII
                                                                                  rsaI                            mseI aflIII
                                        mnlI                    mnlI              csp6I
3901 TCTTTTGTGA GATGTAATTG TTTTTTCAGA GGGCGTGTTG TTTTACCTCA AGTTTTTGTT TTGTACCAAC ACACACACAC ACACACATTC TTAACACATG
     AGAAAACACT CTACATTAAC AAAAAAGTCT CCCGCACAAC AAAATGGAGT TCAAAAACAA AACATGGTTG TGTGTGTGTG TGTGTGTAAG AATTGTGTAC

                                                                                                                  sfuI
                                                                                                                  bstBI
                                                                                                                  bsiCI
                                                                                                                  asuII
                                                                                                        tru9I       ecoRI
                                                                                                        mseI        apoI
                                                                  scfI
4001 TCCTTGTGTG TTTTGAGAGT ATATTATGTA TTTATATTTT GTGCTATCAG ACTGTAGGAT TTGAAGTAGG ACTTTCCTAA ATGTTTAAGA TAAACAGAAT
     AGGAACACAC AAAACTCTCA TATAATACAT AAATATAAAA CACGATAGTC TGACATCCTA AACTTCATCC TGAAAGGATT TACAAATTCT ATTTGTCTTA

     taqI
4101 TC
     AG

length: 4102
```

FIG. 1J

```
wsxfull.6.4.variant      1  MICQKFCVVLLHWEFIYVITAFNLSYP ITPWRFKLSCMPPNSTYDYFLLP
wsxfull.12.1.variant     1  MICQKFCVVLLHWEFIYVITAFNLSYP ITPWRFKLSCMPPNSTYDYFLLP
wsxfull.13.2.variant     1  MICQKFCVVLLHWEFIYVITAFNLSYP ITPWRFKLSCMPPNSTYDYFLLP

wsxfull.6.4.variant     51  AGLSKNTSNSNGHYETAVEPKFNSSGTHFSNLSKTTFHCCFRSEQDRNCS
wsxfull.12.1.variant    51  AGLSKNTSNSNGHYETAVEPKFNSSGTHFSNLSKTTFHCCFRSEQDRNCS
wsxfull.13.2.variant    51  AGLSKNTSNSNGHYETAVEPKFNSSGTHFSNLSKTTFHCCFRSEQDRNCS

wsxfull.6.4.variant    101  LCADNIEGKTFVSTVNSLVFQQIDANWNIQCWLKGDLKLFICYVESLFKN
wsxfull.12.1.variant   101  LCADNIEGKTFVSTVNSLVFQQIDANWNIQCWLKGDLKLFICYVESLFKN
wsxfull.13.2.variant   101  LCADNIEGKTFVSTVNSLVFQQIDANWNIQCWLKGDLKLFICYVESLFKN

wsxfull.6.4.variant    151  LFRNYNYKVHLLYVLPEVLEDSPLVPQKGSFQMVHCNCSVHECCECLVPV
wsxfull.12.1.variant   151  LFRNYNYKVHLLYVLPEVLEDSPLVPQKGSFQMVHCNCSVHECCECLVPV
wsxfull.13.2.variant   151  LFRNYNYKVHLLYVLPEVLEDSPLVPQKGSFQMVHCNCSVHECCECLVPV

wsxfull.6.4.variant    201  PTAKLNDTLLMCLKITSGGVIFQSPLMSVQPINMVKPDPPLGLHMEITDD
wsxfull.12.1.variant   201  PTAKLNDTLLMCLKITSGGVIFQSPLMSVQPINMVKPDPPLGLHMEITDD
wsxfull.13.2.variant   201  PTAKLNDTLLMCLKITSGGVIFQSPLMSVQPINMVKPDPPLGLHMEITDD

wsxfull.6.4.variant    251  GNLKISWSSPPLVPFPLQYQVKYSENSTTVIREADKIVSATSLLVDSILP
wsxfull.12.1.variant   251  GNLKISWSSPPLVPFPLQYQVKYSENSTTVIREADKIVSATSLLVDSILP
wsxfull.13.2.variant   251  GNLKISWSSPPLVPFPLQYQVKYSENSTTVIREADKIVSATSLLVDSILP
```

FIG. 2A

```
wsxfull.6.4.variant    301 GSSYEVQVRGKRLDGPGIWSDWSTPRVFTTQDVIYFPPKILTSVGSNVSF
wsxfull.12.1.variant   301 GSSYEVQVRGKRLDGPGIWSDWSTPRVFTTQDVIYFPPKILTSVGSNVSF
wsxfull.13.2.variant   301 GSSYEVQVRGKRLDGPGIWSDWSTPRVFTTQDVIYFPPKILTSVGSNVSF

wsxfull.6.4.variant    351 HCIYKKENKIVPSKEIVWWMNLAEKIPQSQYDVVSDHVSKVTFFNLNETK
wsxfull.12.1.variant   351 HCIYKKENKIVPSKEIVWWMNLAEKIPQSQYDVVSDHVSKVTFFNLNETK
wsxfull.13.2.variant   351 HCIYKKENKIVPSKEIVWWMNLAEKIPQSQYDVVSDHVSKVTFFNLNETK

wsxfull.6.4.variant    401 PRGKFTYDAVYCCNEHECHHRYAELYVIDVNINISCETDGYLTKMTCRWS
wsxfull.12.1.variant   401 PRGKFTYDAVYCCNEHECHHRYAELYVIDVNINISCETDGYLTKMTCRWS
wsxfull.13.2.variant   401 PRGKFTYDAVYCCNEHECHHRYAELYVIDVNINISCETDGYLTKMTCRWS

wsxfull.6.4.variant    451 TSTIQSLAESTLQLRYHRSSLYCSDIPSIHPISEPKDCYLQSDGFYECIF
wsxfull.12.1.variant   451 TSTIQSLAESTLQLRYHRSSLYCSDIPSIHPISEPKDCYLQSDGFYECIF
wsxfull.13.2.variant   451 TSTIQSLAESTLQLRYHRSSLYCSDIPSIHPISEPKDCYLQSDGFYECIF

wsxfull.6.4.variant    501 QPIFLLSGYTMWIRINHSLGSLDSPPTCVLPDSVVKPLPPSSVKAEITIN
wsxfull.12.1.variant   501 QPIFLLSGYTMWIRINHSLGSLDSPPTCVLPDSVVKPLPPSSVKAEITIN
wsxfull.13.2.variant   501 QPIFLLSGYTMWIRINHSLGSLDSPPTCVLPDSVVKPLPPSSVKAEITIN

wsxfull.6.4.variant    551 IGLLKISWEKPVFPENNLQFQIRYGLSGKEVQWKMYEVYDAKSKSVSLPV
wsxfull.12.1.variant   551 IGLLKISWEKPVFPENNLQFQIRYGLSGKEVQWKMYEVYDAKSKSVSLPV
wsxfull.13.2.variant   551 IGLLKISWEKPVFPENNLQFQIRYGLSGKEVQWKMYEVYDAKSKSVSLPV
```

FIG. 2B

```
wsxfull.6.4.variant     601 PDLCAVYAVQVRCKRLDGLGYWSNWSNPAYTVVMDIKVPMRGPEFWRIIN
wsxfull.12.1.variant    601 PDLCAVYAVQVRCKRLDGLGYWSNWSNPAYTVVMDIKVPMRGPEFWRIIN
wsxfull.13.2.variant    601 PDLCAVYAVQVRCKRLDGLGYWSNWSNPAYTVVMDIKVPMRGPEFWRIIN

wsxfull.6.4.variant     651 GDTMKKEKNVTLLWKPLMKNDSLCSVQRYVINHHTSCNGTWSEDVGNHTK
wsxfull.12.1.variant    651 GDTMKKEKNVTLLWKPLMKNDSLCSVQRYVINHHTSCNGTWSEDVGNHTK
wsxfull.13.2.variant    651 GDTMKKEKNVTLLWKPLMKNDSLCSVQRYVINHHTSCNGTWSEDVGNHTK

wsxfull.6.4.variant     701 FTFLWTEQAHTVTVLAINSIGASVANFNLTFSWPMSKVNIVQSLSAYPLN
wsxfull.12.1.variant    701 FTFLWTEQAHTVTVLAINSIGASVANFNLTFSWPMSKVNIVQSLSAYPLN
wsxfull.13.2.variant    701 FTFLWTEQAHTVTVLAINSIGASVANFNLTFSWPMSKVNIVQSLSAYPLN

wsxfull.6.4.variant     751 SSCVIVSWILSPSDYKLMYFIIEWKNLNEDGEIKWLRISSSVKKYYIHDH
wsxfull.12.1.variant    751 SSCVIVSWILSPSDYKLMYFIIEWKNLNEDGEIKWLRISSSVKKYYIHDH
wsxfull.13.2.variant    751 SSCVIVSWILSPSDYKLMYFIIEWKNLNEDGEIKWLRISSSVKKYYIHDH

                                                          Trans
wsxfull.6.4.variant     801 FIPIEKYQFSLYPIFMEGVGKPKIINSFTQDDIEKHQSDAGLYVIVPVII
wsxfull.12.1.variant    801 FIPIEKYQFSLYPIFMEGVGKPKIINSFTQDDIEKHQSDAGLYVIVPVII
wsxfull.13.2.variant    801 FIPIEKYQFSLYPIFMEGVGKPKIINSFTQDDIEKHQSDAGLYVIVPVII

                    membrane Domain                 Box 1
wsxfull.6.4.variant     851 SSSILLLGTLLISHQRMKKLFWEDVPNPKNCSWAQGLNFQK---------
wsxfull.12.1.variant    851 SSSILLLGTLLISHQRMKKLFWEDVPNPKNCSWAQGLNFQK------MF-
wsxfull.13.2.variant    851 SSSILLLGTLLISHQRMKKLFWEDVPNPKNCSWAQGLNFQKPETFEHLFI
```

FIG. 2C

Box 2

```
wsxfull.13.2.variant  901 KHTASVTCGPLLLEPETISEDISVDTSWKNKDEMMPTTVVSLLSTTDLEK
```

Box 3

```
wsxfull.13.2.variant  951 GSVCISDQFNSVNFSEAEGTEVTYEDESQRQPFVKYATLISNSKPSETGE

wsxfull.6.4.variant   892 ----------------------------------------------R
wsxfull.12.1.variant  894 ------------------------------------------RTPRIVPGH
wsxfull.13.2.variant 1001 EQGLINSSVTKCFSSKNSPLKDSFSNSSWEIEAQAFFILSDQHPNIISPH

wsxfull.6.4.variant   893 ----------------------------------------TDIL-------
wsxfull.12.1.variant  903 ----------------------------------------KDLIF------
wsxfull.13.2.variant 1051 LTFSEGLDELLKLEGNFPEENNDKKSIYYLGVTSIKKRESGVLLTDKSRV

wsxfull.12.1.variant  908 ------------RRCLKAACSLRVITTP--------------------
wsxfull.13.2.variant 1101 SCPFPAPCLFTDIRVLQDSCSHFVENNINLGTSSKKTFASYMPQFQTCST

wsxfull.13.2.variant 1151 QTHKIMENKMCDLTV
```

FIG. 2D

```
wsxfull.6.4.variant      1 GAATTCCGGGTTAAAGCTCTCGTGGCATTATCCTTCAGTGGGGCTATTGG

wsxfull.6.4.variant     51 ACTGACTTTTCTTATGCTGGGATGTGCCTTAGAGGATTATGGATTTGCCA
wsxfull.12.1.variant     1 ------------------------------------GAATTCTCGAGTC
wsxfull.13.2.variant     1 ------------------------------------GAATTCTCGAGTC

wsxfull.6.4.variant    101 GTTCACCCTGACCATCTTGAAAATAAGTTATCTCTGATCTCTGTCTGTAT
wsxfull.12.1.variant    14 GACGGCGGGCGTTAAAGCTCTCGTGGCATTATCCTTCAGTGGGGCTATTG
wsxfull.13.2.variant    14 GACGGCGGGCGTTAAAGCTCTCGTGGCATTATCCTTCAGTGGGGCTATTG

wsxfull.6.4.variant    151 GTTACTTCTCTCCCCTCACCAATGGAGAACAATGTGGGCAAAGTGTACT
wsxfull.12.1.variant    64 GACTGACTTTTCTTATGCTGGGATGTGCCTTAGAGGATTATGGGTGTACT
wsxfull.13.2.variant    64 GACTGACTTTTCTTATGCTGGGATGTGCCTTAGAGGATTATGGGTGTACT

wsxfull.6.4.variant    201 TCTCTGAAGTAAGATGATTTGTCAAAAATTCTGTGTGGTTTTGTTACATT
wsxfull.12.1.variant   114 TCTCTGAAGTAAGATGATTTGTCAAAAATTCTGTGTGGTTTTGTTACATT
wsxfull.13.2.variant   114 TCTCTGAAGTAAGATGATTTGTCAAAAATTCTGTGTGGTTTTGTTACATT

wsxfull.6.4.variant    251 GGGAATTTATTTATGTGATAACTGCGTTTAACTTGTCATATCCAATTACT
wsxfull.12.1.variant   164 GGGAATTTATTTATGTGATAACTGCGTTTAACTTGTCATATCCAATTACT
wsxfull.13.2.variant   164 GGGAATTTATTTATGTGATAACTGCGTTTAACTTGTCATATCCAATTACT

wsxfull.6.4.variant    301 CCTTGGAGATTTAAGTTGTCTTGCATGCCACCAAATTCAACCTATGACTA
wsxfull.12.1.variant   214 CCTTGGAGATTTAAGTTGTCTTGCATGCCACCAAATTCAACCTATGACTA
wsxfull.13.2.variant   214 CCTTGGAGATTTAAGTTGTCTTGCATGCCACCAAATTCAACCTATGACTA
```

FIG. 3A

```
wsxfull.6.4.variant     351 CTTCCTTTTGCCTGCTGGACTCTCAAAGAATACTTCAAATTCGAATGGAC
wsxfull.12.1.variant    264 CTTCCTTTTGCCTGCTGGACTCTCAAAGAATACTTCAAATTCGAATGGAC
wsxfull.13.2.variant    264 CTTCCTTTTGCCTGCTGGACTCTCAAAGAATACTTCAAATTCGAATGGAC

wsxfull.6.4.variant     401 ATTATGAGACAGCTGTTGAACCTAAGTTTAATTCAAGTGGTACTCACTTT
wsxfull.12.1.variant    314 ATTATGAGACAGCTGTTGAACCTAAGTTTAATTCAAGTGGTACTCACTTT
wsxfull.13.2.variant    314 ATTATGAGACAGCTGTTGAACCTAAGTTTAATTCAAGTGGTACTCACTTT

wsxfull.6.4.variant     451 TCTAACTTATCCAAAACAACTTTCCACTGTTGCTTTCGGAGTGAGCAAGA
wsxfull.12.1.variant    364 TCTAACTTATCCAAAACAACTTTCCACTGTTGCTTTCGGAGTGAGCAAGA
wsxfull.13.2.variant    364 TCTAACTTATCCAAAACAACTTTCCACTGTTGCTTTCGGAGTGAGCAAGA

wsxfull.6.4.variant     501 TAGAAACTGCTCCTTATGTGCAGACAACATTGAAGGAAAGACATTTGTTT
wsxfull.12.1.variant    414 TAGAAACTGCTCCTTATGTGCAGACAACATTGAAGGAAAGACATTTGTTT
wsxfull.13.2.variant    414 TAGAAACTGCTCCTTATGTGCAGACAACATTGAAGGAAAGACATTTGTTT

wsxfull.6.4.variant     551 CNACAGTAAATTCTTTAGTTTTTCAACAAATAGATGCAAACTGGAACATA
wsxfull.12.1.variant    464 CAACAGTAAATTCTTTAGTTTTTCAACAAATAGATGCAAACTGGAACATA
wsxfull.13.2.variant    464 CAACAGTAAATTCTTTAGTTTTTCAACAAATAGATGCAAACTGGAACATA

wsxfull.6.4.variant     601 CAGTGCTGGCTAAAAGGAGACTTAAAATTATTCATCTGTTATGTGGAGTC
wsxfull.12.1.variant    514 CAGTGCTGGCTAAAAGGAGACTTAAAATTATTCATCTGTTATGTGGAGTC
wsxfull.13.2.variant    514 CAGTGCTGGCTAAAAGGAGACTTAAAATTATTCATCTGTTATGTGGAGTC
```

FIG. 3B

```
wsxfull.6.4.variant     651 ATTATTTAAGAATCTATTCAGGAATTATAACTATAAGGTCCATCTTTTAT
wsxfull.12.1.variant    564 ATTATTTAAGAATCTATTCAGGAATTATAACTATAAGGTCCATCTTTTAT
wsxfull.13.2.variant    564 ATTATTTAAGAATCTATTCAGGAATTATAACTATAAGGTCCATCTTTTAT

wsxfull.6.4.variant     701 ATGTTCTGCCTGAAGTGTTAGAAGATTCACCTCTGGTTCCCCAAAAAGGC
wsxfull.12.1.variant    614 ATGTTCTGCCTGAAGTGTTAGAAGATTCACCTCTGGTTCCCCAAAAAGGC
wsxfull.13.2.variant    614 ATGTTCTGCCTGAAGTGTTAGAAGATTCACCTCTGGTTCCCCAAAAAGGC

wsxfull.6.4.variant     751 AGTTTTCAGATGGTTCACTGCAATTGCAGTGTTCATGAATGTTGTGAATG
wsxfull.12.1.variant    664 AGTTTTCAGATGGTTCACTGCAATTGCAGTGTTCATGAATGTTGTGAATG
wsxfull.13.2.variant    664 AGTTTTCAGATGGTTCACTGCAATTGCAGTGTTCATGAATGTTGTGAATG

wsxfull.6.4.variant     801 TCTTGTGCCTGTGCCAACAGCCAAACTCAACGACACTCTCCTTATGTGTT
wsxfull.12.1.variant    714 TCTTGTGCCTGTGCCAACAGCCAAACTCAACGACACTCTCCTTATGTGTT
wsxfull.13.2.variant    714 TCTTGTGCCTGTGCCAACAGCCAAACTCAACGACACTCTCCTTATGTGTT

wsxfull.6.4.variant     851 TGAAAATCACATCTGGTGGAGTAATTTTCCAGTCACCTCTAATGTCAGTT
wsxfull.12.1.variant    764 TGAAAATCACATCTGGTGGAGTAATTTTCCAGTCACCTCTAATGTCAGTT
wsxfull.13.2.variant    764 TGAAAATCACATCTGGTGGAGTAATTTTCCAGTCACCTCTAATGTCAGTT

wsxfull.6.4.variant     901 CAGCCCATAAATATGGTGAAGCCTGATCCACCATTAGGTTTGCATATGGA
wsxfull.12.1.variant    814 CAGCCCATAAATATGGTGAAGCCTGATCCACCATTAGGTTTGCATATGGA
wsxfull.13.2.variant    814 CAGCCCATAAATATGGTGAAGCCTGATCCACCATTAGGTTTGCATATGGA
```

FIG. 3C

```
wsxfull.6.4.variant     951  AATCACAGATGATGGTAATTTAAAGATTTCTTGGTCCAGCCCACCATTGG
wsxfull.12.1.variant    864  AATCACAGATGATGGTAATTTAAAGATTTCTTGGTCCAGCCCACCATTGG
wsxfull.13.2.variant    864  AATCACAGATGATGGTAATTTAAAGATTTCTTGGTCCAGCCCACCATTGG

wsxfull.6.4.variant    1001  TACCATTTCCACTTCAATATCAAGTGAAATATTCAGAGAATTCTACAACA
wsxfull.12.1.variant    914  TACCATTTCCACTTCAATATCAAGTGAAATATTCAGAGAATTCTACAACA
wsxfull.13.2.variant    914  TACCATTTCCACTTCAATATCAAGTGAAATATTCAGAGAATTCTACAACA

wsxfull.6.4.variant    1051  GTTATCAGAGAAGCTGACAAGATTGTCTCAGCTACATCCCTGCTAGTAGA
wsxfull.12.1.variant    964  GTTATCAGAGAAGCTGACAAGATTGTCTCAGCTACATCCCTGCTAGTAGA
wsxfull.13.2.variant    964  GTTATCAGAGAAGCTGACAAGATTGTCTCAGCTACATCCCTGCTAGTAGA

wsxfull.6.4.variant    1101  CAGTATACTTCCTGGGTCTTCGTATGAGGTTCAGGTGAGGGGCAAGAGAC
wsxfull.12.1.variant   1014  CAGTATACTTCCTGGGTCTTCGTATGAGGTTCAGGTGAGGGGCAAGAGAC
wsxfull.13.2.variant   1014  CAGTATACTTCCTGGGTCTTCGTATGAGGTTCAGGTGAGGGGCAAGAGAC

wsxfull.6.4.variant    1151  TGGATGGCCCAGGAATCTGGAGTGACTGGAGTACTCCTCGTGTCTTTACC
wsxfull.12.1.variant   1064  TGGATGGCCCAGGAATCTGGAGTGACTGGAGTACTCCTCGTGTCTTTACC
wsxfull.13.2.variant   1064  TGGATGGCCCAGGAATCTGGAGTGACTGGAGTACTCCTCGTGTCTTTACC

wsxfull.6.4.variant    1201  ACACAAGATGTCATATACTTTCCACCTAAAATTCTGACAAGTGTTGGGTC
wsxfull.12.1.variant   1114  ACACAAGATGTCATATACTTTCCACCTAAAATTCTGACAAGTGTTGGGTC
wsxfull.13.2.variant   1114  ACACAAGATGTCATATACTTTCCACCTAAAATTCTGACAAGTGTTGGGTC
```

FIG. 3D

```
wsxfull.6.4.variant    1251 TAATGTTTCTTTTCACTGCATCTATAAGAAGGAAAACAAGATTGTTCCCT
wsxfull.12.1.variant   1164 TAATGTTTCTTTTCACTGCATCTATAAGAAGGAAAACAAGATTGTTCCCT
wsxfull.13.2.variant   1164 TAATGTTTCTTTTCACTGCATCTATAAGAAGGAAAACAAGATTGTTCCCT

wsxfull.6.4.variant    1301 CAAAAGAGATTGTTTGGTGGATGAATTTAGCTGAGAAAATTCCTCAAAGC
wsxfull.12.1.variant   1214 CAAAAGAGATTGTTTGGTGGATGAATTTAGCTGAGAAAATTCCTCAAAGC
wsxfull.13.2.variant   1214 CAAAAGAGATTGTTTGGTGGATGAATTTAGCTGAGAAAATTCCTCAAAGC

wsxfull.6.4.variant    1351 CAGTATGATGTTGTGAGTGATCATGTTAGCAAAGTTACTTTTTTCAATCT
wsxfull.12.1.variant   1264 CAGTATGATGTTGTGAGTGATCATGTTAGCAAAGTTACTTTTTTCAATCT
wsxfull.13.2.variant   1264 CAGTATGATGTTGTGAGTGATCATGTTAGCAAAGTTACTTTTTTCAATCT

wsxfull.6.4.variant    1401 GAATGAAACCAAACCTCGAGGAAAGTTTACCTATGATGCAGTGTACTGCT
wsxfull.12.1.variant   1314 GAATGAAACCAAACCTCGAGGAAAGTTTACCTATGATGCAGTGTACTGCT
wsxfull.13.2.variant   1314 GAATGAAACCAAACCTCGAGGAAAGTTTACCTATGATGCAGTGTACTGCT

wsxfull.6.4.variant    1451 GCAATGAACATGAATGCCATCATCGCTATGCTGAATTATATGTGATTGAT
wsxfull.12.1.variant   1364 GCAATGAACATGAATGCCATCATCGCTATGCTGAATTATATGTGATTGAT
wsxfull.13.2.variant   1364 GCAATGAACATGAATGCCATCATCGCTATGCTGAATTATATGTGATTGAT

wsxfull.6.4.variant    1501 GTCAATATCAATATCTCATGTGAAACTGATGGGTACTTAACTAAAATGAC
wsxfull.12.1.variant   1414 GTCAATATCAATATCTCATGTGAAACTGATGGGTACTTAACTAAAATGAC
wsxfull.13.2.variant   1414 GTCAATATCAATATCTCATGTGAAACTGATGGGTACTTAACTAAAATGAC
```

FIG. 3E

```
wsxfull.6.4.variant    1551 TTGCAGATGGTCAACCAGTACAATCCAGTCACTTGCGGAAAGCACTTTGC
wsxfull.12.1.variant   1464 TTGCAGATGGTCAACCAGTACAATCCAGTCACTTGCGGAAAGCACTTTGC
wsxfull.13.2.variant   1464 TTGCAGATGGTCAACCAGTACAATCCAGTCACTTGCGGAAAGCACTTTGC

wsxfull.6.4.variant    1601 AATTGAGGTATCATAGGAGCAGCCTTTACTGTTCTGATATTCCATCTATT
wsxfull.12.1.variant   1514 AATTGAGGTATCATAGGAGCAGCCTTTACTGTTCTGATATTCCATCTATT
wsxfull.13.2.variant   1514 AATTGAGGTATCATAGGAGCAGCCTTTACTGTTCTGATATTCCATCTATT

wsxfull.6.4.variant    1651 CATCCCATATCTGAGCCCAAAGATTGCTATTTGCAGAGTGATGGTTTTTA
wsxfull.12.1.variant   1564 CATCCCATATCTGAGCCCAAAGATTGCTATTTGCAGAGTGATGGTTTTTA
wsxfull.13.2.variant   1564 CATCCCATATCTGAGCCCAAAGATTGCTATTTGCAGAGTGATGGTTTTTA

wsxfull.6.4.variant    1701 TGAATGCATTTTCCAGCCAATCTTCCTATTATCTGGCTACACAATGTGGA
wsxfull.12.1.variant   1614 TGAATGCATTTTCCAGCCAATCTTCCTATTATCTGGCTACACAATGTGGA
wsxfull.13.2.variant   1614 TGAATGCATTTTCCAGCCAATCTTCCTATTATCTGGCTACACAATGTGGA

wsxfull.6.4.variant    1751 TTAGGATCAATCACTCTCTAGGTTCACTTGACTCTCCACCAACATGTGTC
wsxfull.12.1.variant   1664 TTAGGATCAATCACTCTCTAGGTTCACTTGACTCTCCACCAACATGTGTC
wsxfull.13.2.variant   1664 TTAGGATCAATCACTCTCTAGGTTCACTTGACTCTCCACCAACATGTGTC

wsxfull.6.4.variant    1801 CTTCCTGATTCTGTGGTGAAGCCACTGCCTCCATCCAGTGTGAAAGCAGA
wsxfull.12.1.variant   1714 CTTCCTGATTCTGTGGTGAAGCCACTGCCTCCATCCAGTGTGAAAGCAGA
wsxfull.13.2.variant   1714 CTTCCTGATTCTGTGGTGAAGCCACTGCCTCCATCCAGTGTGAAAGCAGA
```

FIG. 3F

```
wsxfull.6.4.variant    1851 AATTACTATAAACATTGGATTATTGAAAATATCTTGGGAAAAGCCAGTCT
wsxfull.12.1.variant   1764 AATTACTATAAACATTGGATTATTGAAAATATCTTGGGAAAAGCCAGTCT
wsxfull.13.2.variant   1764 AATTACTATAAACATTGGATTATTGAAAATATCTTGGGAAAAGCCAGTCT

wsxfull.6.4.variant    1901 TTCCAGAGAATAACCTTCAATTCCAGATTCGCTATGGTTTAAGTGGAAAA
wsxfull.12.1.variant   1814 TTCCAGAGAATAACCTTCAATTCCAGATTCGCTATGGTTTAAGTGGAAAA
wsxfull.13.2.variant   1814 TTCCAGAGAATAACCTTCAATTCCAGATTCGCTATGGTTTAAGTGGAAAA

wsxfull.6.4.variant    1951 GAAGTACAATGGAAGATGTATGAGGTTTATGATGCAAAATCAAAATCTGT
wsxfull.12.1.variant   1864 GAAGTACAATGGAAGATGTATGAGGTTTATGATGCAAAATCAAAATCTGT
wsxfull.13.2.variant   1864 GAAGTACAATGGAAGATGTATGAGGTTTATGATGCAAAATCAAAATCTGT

wsxfull.6.4.variant    2001 CAGTCTCCCAGTTCCAGACTTGTGTGCAGTCTATGCTGTTCAGGTGCGCT
wsxfull.12.1.variant   1914 CAGTCTCCCAGTTCCAGACTTGTGTGCAGTCTATGCTGTTCAGGTGCGCT
wsxfull.13.2.variant   1914 CAGTCTCCCAGTTCCAGACTTGTGTGCAGTCTATGCTGTTCAGGTGCGCT

wsxfull.6.4.variant    2051 GTAAGAGGCTAGATGGACTGGGATATTGGAGTAATTGGAGCAATCCAGCC
wsxfull.12.1.variant   1964 GTAAGAGGCTAGATGGACTGGGATATTGGAGTAATTGGAGCAATCCAGCC
wsxfull.13.2.variant   1964 GTAAGAGGCTAGATGGACTGGGATATTGGAGTAATTGGAGCAATCCAGCC

wsxfull.6.4.variant    2101 TACACAGTTGTCATGGATATAAAAGTTCCTATGAGAGGACCTGAATTTTG
wsxfull.12.1.variant   2014 TACACAGTTGTCATGGATATAAAAGTTCCTATGAGAGGACCTGAATTTTG
wsxfull.13.2.variant   2014 TACACAGTTGTCATGGATATAAAAGTTCCTATGAGAGGACCTGAATTTTG
```

FIG. 3G

```
wsxfull.6.4.variant    2151 GAGAATAATTAATGGAGATACTATGAAAAAGGAGAAAAATGTCACTTTAC
wsxfull.12.1.variant   2064 GAGAATAATTAATGGAGATACTATGAAAAAGGAGAAAAATGTCACTTTAC
wsxfull.13.2.variant   2064 GAGAATAATTAATGGAGATACTATGAAAAAGGAGAAAAATGTCACTTTAC

wsxfull.6.4.variant    2201 TTTGGAAGCCCCTGATGAAAAATGACTCATTGTGCAGTGTTCAGAGATAT
wsxfull.12.1.variant   2114 TTTGGAAGCCCCTGATGAAAAATGACTCATTGTGCAGTGTTCAGAGATAT
wsxfull.13.2.variant   2114 TTTGGAAGCCCCTGATGAAAAATGACTCATTGTGCAGTGTTCAGAGATAT

wsxfull.6.4.variant    2251 GTGATAAACCATCATACTTCCTGCAATGGAACATGGTCAGAAGATGTGGG
wsxfull.12.1.variant   2164 GTGATAAACCATCATACTTCCTGCAATGGAACATGGTCAGAAGATGTGGG
wsxfull.13.2.variant   2164 GTGATAAACCATCATACTTCCTGCAATGGAACATGGTCAGAAGATGTGGG

wsxfull.6.4.variant    2301 AAATCACACGAAATTCACTTTCCTGTGGACAGAGCAAGCACATACTGTTA
wsxfull.12.1.variant   2214 AAATCACACGAAATTCACTTTCCTGTGGACAGAGCAAGCACATACTGTTA
wsxfull.13.2.variant   2214 AAATCACACGAAATTCACTTTCCTGTGGACAGAGCAAGCACATACTGTTA

wsxfull.6.4.variant    2351 CGGTTCTGGCCATCAATTCAATTGGTGCTTCTGTTGCAAATTTTAATTTA
wsxfull.12.1.variant   2264 CGGTTCTGGCCATCAATTCAATTGGTGCTTCTGTTGCAAATTTTAATTTA
wsxfull.13.2.variant   2264 CGGTTCTGGCCATCAATTCAATTGGTGCTTCTGTTGCAAATTTTAATTTA

wsxfull.6.4.variant    2401 ACCTTTTCATGGCCTATGAGCAAAGTAAATATCGTGCAGTCACTCAGTGC
wsxfull.12.1.variant   2314 ACCTTTTCATGGCCTATGAGCAAAGTAAATATCGTGCAGTCACTCAGTGC
wsxfull.13.2.variant   2314 ACCTTTTCATGGCCTATGAGCAAAGTAAATATCGTGCAGTCACTCAGTGC
```

FIG. 3H

```
wsxfull.6.4.variant    2451 TTATCCTTTAAACAGCAGTTGTGTGATTGTTTCCTGGATACTATCACCCA
wsxfull.12.1.variant   2364 TTATCCTTTAAACAGCAGTTGTGTGATTGTTTCCTGGATACTATCACCCA
wsxfull.13.2.variant   2364 TTATCCTTTAAACAGCAGTTGTGTGATTGTTTCCTGGATACTATCACCCA

wsxfull.6.4.variant    2501 GTGATTACAAGCTAATGTATTTTATTATTGAGTGGAAAAATCTTAATGAA
wsxfull.12.1.variant   2414 GTGATTACAAGCTAATGTATTTTATTATTGAGTGGAAAAATCTTAATGAA
wsxfull.13.2.variant   2414 GTGATTACAAGCTAATGTATTTTATTATTGAGTGGAAAAATCTTAATGAA

wsxfull.6.4.variant    2551 GATGGTGAAATAAAATGGCTTAGAATCTCTTCATCTGTTAAGAAGTATTA
wsxfull.12.1.variant   2464 GATGGTGAAATAAAATGGCTTAGAATCTCTTCATCTGTTAAGAAGTATTA
wsxfull.13.2.variant   2464 GATGGTGAAATAAAATGGCTTAGAATCTCTTCATCTGTTAAGAAGTATTA

wsxfull.6.4.variant    2601 TATCCATGATCATTTTATCCCCATTGAGAAGTACCAGTTCAGTCTTTACC
wsxfull.12.1.variant   2514 TATCCATGATCATTTTATCCCCATTGAGAAGTACCAGTTCAGTCTTTACC
wsxfull.13.2.variant   2514 TATCCATGATCATTTTATCCCCATTGAGAAGTACCAGTTCAGTCTTTACC

wsxfull.6.4.variant    2651 CAATATTTATGGAAGGAGTGGGAAAACCAAAGATAATTAATAGTTTCACT
wsxfull.12.1.variant   2564 CAATATTTATGGAAGGAGTGGGAAAACCAAAGATAATTAATAGTTTCACT
wsxfull.13.2.variant   2564 CAATATTTATGGAAGGAGTGGGAAAACCAAAGATAATTAATAGTTTCACT

wsxfull.6.4.variant    2701 CAAGATGATATTGAAAAACACCAGAGTGATGCAGGTTTATATGTAATTGT
wsxfull.12.1.variant   2614 CAAGATGATATTGAAAAACACCAGAGTGATGCAGGTTTATATGTAATTGT
wsxfull.13.2.variant   2614 CAAGATGATATTGAAAAACACCAGAGTGATGCAGGTTTATATGTAATTGT
```

FIG. 3I

```
wsxfull.6.4.variant    2751 GCCAGTAATTATTTCCTCTTCCATCTTATTGCTTGGAACATTATTAATAT
wsxfull.12.1.variant   2664 GCCAGTAATTATTTCCTCTTCCATCTTATTGCTTGGAACATTATTAATAT
wsxfull.13.2.variant   2664 GCCAGTAATTATTTCCTCTTCCATCTTATTGCTTGGAACATTATTAATAT

wsxfull.6.4.variant    2801 CACACCAAAGAATGAAAAAGCTATTTTGGGAAGATGTTCCGAACCCCAAG
wsxfull.12.1.variant   2714 CACACCAAAGAATGAAAAAGCTATTTTGGGAAGATGTTCCGAACCCCAAG
wsxfull.13.2.variant   2714 CACACCAAAGAATGAAAAAGCTATTTTGGGAAGATGTTCCGAACCCCAAG

wsxfull.6.4.variant    2851 AATTGTTCCTGGGCACAAGGACTTAATTTTCAGAAGAGACGGACATTCT
wsxfull.12.1.variant   2764 AATTGTTCCTGGGCACAAGGACTTAATTTTCAGAAGATGTTCCGAACCCC
wsxfull.13.2.variant   2764 AATTGTTCCTGGGCACAAGGACTTAATTTTCAGAAGCCAGAAACGTTTGA

wsxfull.6.4.variant    2901 TTGAAGTCTAATCATGATCACTACAGATGAACCCAATGTGCCAACTTCCC
wsxfull.12.1.variant   2814 AAGAATTGTTCCTGGGCACAAGGACTTAATTTTCAGAAGATGCTTGAAGG
wsxfull.13.2.variant   2814 GCATCTTTTATCAAGCATACAGCATCAGTGACATGTGGTCCTCTTCTTT

wsxfull.6.4.variant    2951 AACAGTCTATAGAGTATTAGAAGATTTTACATTTTGAAGAAGGGCCGGA
wsxfull.12.1.variant   2864 CAGCATGTTCGTTAAGAGTCATCACCACTCCCTAATCTCAAGTACCCAGG
wsxfull.13.2.variant   2864 TGGAGCCTGAACAATTTCAGAAGATATCAGTGTTGATACATCATGGAAA

wsxfull.6.4.variant    3001 ATTC.....................................................
wsxfull.12.1.variant   2914 GACACAAACACTGCGGAAGGCCACAGGGTCCTCTGCATAGGAAAACCAGA
wsxfull.13.2.variant   2914 AATAAAGATGAGATGATGCCAACAACTGTGGTCTCTCTACTTTCAACAAC
```

FIG. 3J

```
wsxfull.12.1.variant 2964 GACCTTTGTTCACTTGTTTATCTGCTGACCCTCCCTCCACTATTGTCCTA
wsxfull.13.2.variant 2964 AGATCTTGAAAAGGGTTCTGTTTGTATTAGTGACCAGTTCAACAGTGTTA

wsxfull.12.1.variant 3014 TGACCCTGCCAAATCCCCCCTCTGTGAGAACACCCAAGAATGATCAATAA
wsxfull.13.2.variant 3014 ACTTCTCTGAGGCTGAGGGTACTGAGGTAACCTATGAGGACGAAGCCAG

wsxfull.12.1.variant 3064 AAAAAAAAAAAAAAAAAAAAGTCGACTCGAGAATTC.............
wsxfull.13.2.variant 3064 AGACAACCCTTTGTTAAATACGCCACGCTGATCAGCAACTCTAAACCAAG

wsxfull.13.2.variant 3114 TGAAACTGGTGAAGAACAAGGGCTTATAAATAGTTCAGTCACCAAGTGCT

wsxfull.13.2.variant 3164 TCTCTAGCAAAAATTCTCCGTTGAAGGATTCTTTCTCTAATAGCTCATGG

wsxfull.13.2.variant 3214 GAGATAGAGGCCCAGGCATTTTTTATATTATCAGATCAGCATCCCAACAT

wsxfull.13.2.variant 3264 AATTTCACCACACCTCACATTCTCAGAAGGATTGGATGAACTTTTGAAAT

wsxfull.13.2.variant 3314 TGGAGGGAAATTTCCCTGAAGAAAATAATGATAAAAAGTCTATCTATTAT

wsxfull.13.2.variant 3364 TTAGGGGTCACCTCAATCAAAAAGAGAGAGAGTGGTGTGCTTTTGACTGA

wsxfull.13.2.variant 3414 CAAGTCAAGGGTATCGTGCCCATTCCCAGCCCCCTGTTTATTCACGGACA

wsxfull.13.2.variant 3464 TCAGAGTTCTCCAGGACAGTTGCTCACACTTTGTAGAAAATAATATCAAC
```

FIG. 3K

```
wsxfull.13.2.variant 3514 T T A G G A A C T T C T A G T A A G A A G A C T T T T G C A T C T T A C A T G C C T C A A T T C C A

wsxfull.13.2.variant 3564 A A C T T G T T C T A C T C A G A C T C A T A A G A T C A T G G A A A A C A A G A T G T G T G A C C

wsxfull.13.2.variant 3614 T A A C T G T G T A A T T T C A C T G A A G A A A C C T T C A G A T T T G T G T T A T A A T G G G T

wsxfull.13.2.variant 3664 A A T A T A A A G T G T A A T A G A T T A T A G T T G T G G G T G G G A G A G A G A A A A G A A A C

wsxfull.13.2.variant 3714 C A G A G T C A A A T T T G A A A A T A A T T G T T C C A A A T G A A T G T T G T C T G T T T G T T

wsxfull.13.2.variant 3764 C T C T C T T A G T A A C A T A G A C A A A A A A T T T G A G A A A G C C T T C A T A A G C C T A C

wsxfull.13.2.variant 3814 C A A T G T A G A C A C G C T C T T C T A T T T T A T T C C C A A G C T C T A G T G G G A A G G T C

wsxfull.13.2.variant 3864 C C T T G T T T C C A G C T A G A A A T A A G C C C A A C A G A C A C C A T C T T T T G T G A G A T

wsxfull.13.2.variant 3914 G T A A T T G T T T T T T C A G A G G G C G T G T T G T T T T A C C T C A A G T T T T T G T T T T G

wsxfull.13.2.variant 3964 T A C C A A C A C A C A C A C A C A C A C A C A T T C T T A A C A C A T G T C C T T G T G T G T T T

wsxfull.13.2.variant 4014 T G A G A G T A T A T T A T G T A T T T A T A T T T T G T G C T A T C A G A C T G T A G G A T T T G

wsxfull.13.2.variant 4064 A A G T A G G A C T T T C C T A A A T G T T T A A G A T A A A C A G A A T T C
```

FIG. 3L

```
wsxfull.13.2.variant   1  MICQKFCVVLLHWEFIYVITAFNLSYPITPWRFKLSCMPPNSTYDYFLLP
mu.wsx.ecd             1  MMCQKFYVVLLHWEFLYVIAALNLAYPISPWKFKLFCGPPNTTDDSFLSP

wsxfull.13.2.variant   51 AGLSKNTSNSNGHYETAVEPKFNSSGTHFSNLSKTTFHCCFRSEQDRNCS
mu.wsx.ecd             51 AGAPNNASALKGASEAIVEAKFNSSGIYVPELSKTVFHCCFGNEQGQNCS

wsxfull.13.2.variant   101 LCADNIEGKTFVSTVNSLVFQQIDANWNIQCWLKGDLKLFICYVESLFKN
mu.wsx.ecd             101 ALTDNTEGKTLASVVKASVFRQLGVNWDIECWMKGDLTLFICHMEPLPKN

wsxfull.13.2.variant   151 LFRNYNYKVHLLYVLPEVLEDSPLVPQKGSFQMVHCNCSVHECCECLVPV
mu.wsx.ecd             151 PFKNYDSKVHLLYDLPEVIDDSPLPPLKDSFQTVQCNCSLRG-CECHVPV

wsxfull.13.2.variant   201 PTAKLNDTLLMCLKITSGGVIFQSPLMSVQPINMVKPDPPLGLHMEITDD
mu.wsx.ecd             200 PRAKLNYALLMYLEITSAGVSFQSPLMSLQPMLVVKPDPPLGLHMEVTDD

wsxfull.13.2.variant   251 GNLKISWSSPPLVPFPLQYQVKYSENSTTVIREADKIVSATSLLVDSILP
mu.wsx.ecd             250 GNLKISWDSQTMAPFPLQYQVKYLENS-TIVREAAEIVSATSLLVDSVLP
```

FIG. 4A

```
wsxfull.13.2.variant  301 GSSYEVQVRGKRLDGPGIWSDWSTPRVFTTQDVIYFPPKILTSVGSNVSF
mu.wsx.ecd            299 GSSYEVQVRSKRLDGSGVWSDWSSPQVFTTQDVVYFPPKILTSVGSNASF

wsxfull.13.2.variant  351 HCIYKKENKIVPSKEIVWWMNLAEKIPQSQYDVVSDHVSKVTFFNLNETK
mu.wsx.ecd            349 HCIYKNENQIVSSKQIVWWRNLAEKIPEIQYSIVSDRVSKVTFSNLKATR

wsxfull.13.2.variant  401 PRGKFTYDAVYCCNEHECHHRYAELYVIDVNINISCETDGYLTKMTCRWS
mu.wsx.ecd            399 PRGKFTYDAVYCCNEQACHHRYAELYVIDVNINISCETDGYLTKMTCRWS

wsxfull.13.2.variant  451 TSTIQSLAESTLQLRYHRSSLYCSDIPSIHPISEPKDCYLQSDGFYECIF
mu.wsx.ecd            449 PSTIQSLVGSTVQLRYHRCSLYCPDSPSIHPTSEPKTASYRETAFMNVFS

wsxfull.13.2.variant  501 QPIFLLSGYTMWIRINHSLGSLDSPPTCVLPDSVVKPLPPSSVKAEITIN
mu.wsx.ecd            499 SQSFYYLAIQCGFRINHSLGSLDSPPTCVLPDSVVKPLPPSNVKAEITVN

wsxfull.13.2.variant  551 IGLLKISWEKPVFPENNLQFQIRYGLSGKEVQWKMYEVYDAKSKSVSLPV
mu.wsx.ecd            549 TGLLKVSWEKPVFPENNLQFQIRYGLSGKEIQWKTHEVFDAKSKSASLLV
```

FIG. 4B

```
wsxfull.13.2.variant  601 P D L C A V Y A V Q V R C K R L D G L G Y W S N W S N P A Y T V V M D I K V P M R G P E F W R I I N
mu.wsx.ecd            599 S D L C A V Y V V Q V R C R R L D G L G Y W S N W S S P A Y T L V M D V K V P M R G P E F W R K M D

wsxfull.13.2.variant  651 G D T M K K E K N V T L L W K P L M K N D S L C S V Q R Y V I N H H T S C N G T W S E D V G N H T K
mu.wsx.ecd            649 G D V T K K E R N V T L L W K P L T K N D S L C S V R R Y V V K H R T A H N G T W S E D V G N R T N

wsxfull.13.2.variant  701 F T F L W T E Q A H T V T V L A I N S I G A S V A N F N L T F S W P M S K V N I V Q S L S A Y P L N
mu.wsx.ecd            699 L T F L W T E P A H T V T V L A V N S L G A S L V N F N L T F S W P M S K V S A V E S L S A Y P L S

wsxfull.13.2.variant  751 S S C V I V S W I L S P S D Y K L M Y F I I E W K N L N E D G E I K W L R I S S S V K K Y Y I H D H
mu.wsx.ecd            749 S S C V I L S W T L S P D D Y S L L Y L V I E W K I L N E D D G M K W

wsxfull.13.2.variant  801 F I P I E K Y Q F S L Y P I F M E G V G K P K I I N S F T Q D D I E K H Q S D A G L Y V I V P V I I

wsxfull.13.2.variant  851 S S S I L L L G T L L I S H Q R M K K L F W E D V P N P K N C S W A Q G L N F Q K P E T F E H L F I
```

FIG. 4C

```
wsxfull.13.2.variant  901 K H T A S V T C G P L L L E P E T I S E D I S V D T S W K N K D E M M P T T V V S L L S T T D L E K

wsxfull.13.2.variant  951 G S V C I S D Q F N S V N F S E A E G T E V T Y E D E S Q R Q P F V K Y A T L I S N S K P S E T G E

wsxfull.13.2.variant 1001 E Q G L I N S S V T K C F S S K N S P L K D S F S N S S W E I E A Q A F F I L S D Q H P N I I S P H

wsxfull.13.2.variant 1051 L T F S E G L D E L L K L E G N F P E E N N D K K S I Y Y L G V T S I K K R E S G V L L T D K S R V

wsxfull.13.2.variant 1101 S C P F P A P C L F T D I R V L Q D S C S H F V E N N I N L G T S S K K T F A S Y M P Q F Q T C S T

wsxfull.13.2.variant 1151 Q T H K I M E N K M C D L T V
```

FIG. 4D

```
mu.wsx.ecd                1 G G G C C C C C C C C T C G A A G T C G A C G G T A T C G A T A A G C T T G A T A T C G A A T T C C G

mu.wsx.ecd               51 G C C G G G A C A C A G G T G G G A C A C T C T T T T A G T C C T C A A T C C C T G G C G C G A G G

mu.wsx.ecd              101 C C A C C C A A G G C A A C G C A G G A C G C A G G G C G T T T G G G G A C C A G G C A G C A G A C

mu.wsx.ecd              151 T G G G G C G G T A C C T G C G G A G A G C C A C G C A A C T T C T C C A G G C C T C T G A C T A C

mu.wsx.ecd              201 T T T G G A A A C T G C C C G G G G C T G C G A C A T C A A C C C C T T A A G T C C C G G A G G C G

mu.wsx.ecd              251 G A A A G A G G G T G G G T T G G T T T G A A A G A C A C A A G G A A G A A A A A T G T G C T G T G

mu.wsx.ecd              301 G G G C G G G T T A A G T T T C C C A C C C T C T T C C C C C T T C C C G A G C A A A T T A G A A A

mu.wsx.ecd              351 C A A A A C A A A T A G A A A A G C C A G C C C T C C G G C C A A C C A A A G C C C C A A G C G G A
wsxfull.13.2.variant      1 . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . G A A T T C T C G A G T C G A C
```

FIG. 5A

```
mu.wsx.ecd              401 GCCCCAAGCGGAGCCCCAGCCGGAGCACTCCTTTAAAAGGATTTGCAGCG
wsxfull.13.2.variant     17 GGCGGGCGTTAAAGCTCTCGTGGCATTATCCTTCAGTGGGGCT---ATTG

mu.wsx.ecd              451 GTGAGGAAAAAACCAGACCCGACCGAGGAATCGTTCTGCAAATCCAGGTG
wsxfull.13.2.variant     64 GACTGACTTTTCTTATGCTGGGATGTG----CCTTAGAGGATTATGGGTG

mu.wsx.ecd              501 TACACCTCTGAAGAAAGATGATGTGTCAGAAATTCTATGTGGTTTTGTTA
wsxfull.13.2.variant    110 TACTTCTCTGAAGTAAGATGATTGTCAAAAATTCTGTGTGGTTTTGTTA

mu.wsx.ecd              551 CACTGGGAATTTCTTTATGTGATAGCTGCACTTAACCTGGCATATCCAAT
wsxfull.13.2.variant    160 CATTGGGAATTTATTTATGTGATAACTGCGTTTAACTTGTCATATCCAAT

mu.wsx.ecd              601 CTCTCCCTGGAAATTTTAAGTTGTTTTGTGGACCACCGAACACAACCGGATG
wsxfull.13.2.variant    210 TACTCCTTGGAGATTTTAAGTTGTCTTGCATGCCACCAAATTCAACCTATG

mu.wsx.ecd              651 ACTCCTTTCTCTCAAACCTGCTGGAGCCCCAAACAATGCCTCGGCTTTGAAG
wsxfull.13.2.variant    260 ACTACTTCCTTTGCCTGCTGGACTCTCAAAGAATACTTCAAATTCGAAT
```

FIG. 5B

```
mu.wsx.ecd          701 GGGGCTTCTGAAGCAATTGTTGAAGCTAAATTTAATTCAAGTGGTATCTA
wsxfull.13.2.variant 310 GGACATTATGAGACAGCTGTTGAACCTAAGTTTAATTCAAGTGGTACTCA

mu.wsx.ecd          751 CGTTCCTGAGTTATCCAAAACAGTCTTCCACTGTTGCTTTGGGAATGAGC
wsxfull.13.2.variant 360 CTTTCTAACTTATCCAAAACAACTTTCCACTGTTGCTTTCGGAGTGAGC

mu.wsx.ecd          801 AAGGTCAAAACTGCTCTGCACTCACAGACAACACTGAAGGGAAGACACTG
wsxfull.13.2.variant 410 AAGATAGAAACTGCTCCTTATGTGCAGACAACATTGAAGGAAAGACATTT

mu.wsx.ecd          851 GCTTCAGTAGTGAAGGCTTCAGTTTTTCGCCAGCTAGGTGTAAACTGGGA
wsxfull.13.2.variant 460 GTTTCAACAGTAAATTCTTTAGTTTTTCAACAATAGATGCAAACTGGAA

mu.wsx.ecd          901 CATAGAGTGCTGGATGAAAGGGGACTTGACATTATTCATCTGTCATATGG
wsxfull.13.2.variant 510 CATACAGTGCTGGCTAAAAGGAGACTTAAAATTATTCATCTGTTATGTGG

mu.wsx.ecd          951 AGCCATTACCTAAGAACCCCTTCAAGAATTATGACTCTAAGGTCCATCTT
wsxfull.13.2.variant 560 AGTCATTATTAAGAATCCTATTCAGGAATTATAACTATAAGGTCCATCTT
```

FIG. 5C

```
mu.wsx.ecd          1001 TTATATGATCTGCCTGAAGTCATAGATGATTCGCCTCTGCCCCCACTGAA
wsxfull.13.2.variant 610 TTATATGTCTGCCTGAAGTGTAGAAGATTCACCTCTGGTTCCCCAAAA

mu.wsx.ecd          1051 AGACAGCTTTCAGACTGTCCAATGCAACTGCAGTCTTCGGG---GATGTG
wsxfull.13.2.variant 660 AGGCAGTTTTCAGATGGTTCACTGCAATTGCAGTGTTCATGAATGTTGTG

mu.wsx.ecd          1098 AATGTCATGTGCCAGTACCCAGAGCCAAACTCAACTACGCTCTTCTGATG
wsxfull.13.2.variant 710 AATGTCTTGTGCCTGTGCCAACAGCCAAACTCAACGACACTCTCCTTATG

mu.wsx.ecd          1148 TATTTGGAAATCACATCTGCCGGTGTGAGTTTTCAGTCACCTCTGATGTC
wsxfull.13.2.variant 760 TGTTTGAAAATCACATCTGGTGGAGTAATTTCAGTCACCTCTAATGTC

mu.wsx.ecd          1198 ACTGCAGCCCATGCTTGTTGTGAAACCCGATCCACCCTTAGGTTTGCATA
wsxfull.13.2.variant 810 AGTTCAGCCCATAAATATGGTGAAGCCTGATCCACCATTAGGTTTGCATA

mu.wsx.ecd          1248 TGGAAGTCACAGATGATGGTAATTTAAAGATTTCTTGGGACAGCCAAACA
wsxfull.13.2.variant 860 TGGAAATCACAGATGATGGTAATTTAAAGATTTCTTGGTCCAGCCCACCA
```

FIG. 5D

```
mu.wsx.ecd          1298 ATGGCACCATTTCCGCTTCAATATCAGGTGAAAATATTTAGAGAATTCTAC
wsxfull.13.2.variant 910 TTGGTACCATTTCCACTTCAATATCAAGTGAAAATATTCAGAGAATTCTAC

mu.wsx.ecd          1348 AA...TTGTAAGAGAGGCTGCTGAAATTGTCTCAGCTACATCTCTGCTGG
wsxfull.13.2.variant 960 AACAGTTATCAGAGAAGCTGACAAGATTGTCTCAGCTACATCCCTGCTAG

mu.wsx.ecd          1395 TAGACAGTGTGCTTCCTGGATCTTCATATGAGGTCCAGGTGAGGAGCAAG
wsxfull.13.2.variant 1010 TAGACAGTATACTTCCTGGGTCTTCGTATGAGGTTCAGGTGAGGGCAAG

mu.wsx.ecd          1445 AGACTGGATGGTTCAGGAGTCTGGAGTGACTGGAGTTCACCTCAAGTCTT
wsxfull.13.2.variant 1060 AGACTGGATGGCCCAGGAATCTGGAGTGACTGGAGTACTCCTCGTGTCTT

mu.wsx.ecd          1495 TACCACACAAGATGTTGTGTATTTTCCACCCAAAATTCTGACTAGTGTTG
wsxfull.13.2.variant 1110 TACCACACAAGATGTCATATACTTTCCACCTAAAATTCTGACAAGTGTTG

mu.wsx.ecd          1545 GATCGAATGCTTCCTTTCATTGCATCTACAAAACGAAAACCAGATTGTC
wsxfull.13.2.variant 1160 GGTCTAATGTTTCTTTCACTGCATCTATAAGAAGGAAAACAAGATTGTT
```

FIG. 5E

```
mu.wsx.ecd                1595 TCCTCAAAACAGATAGTTTGGTGGAGGAATCTAGCTGAGAAAATCCCTGA
wsxfull.13.2.variant      1210 CCCTCAAAAGAGATTGTTTGGTGGATGAATTTAGCTGAGAAAATTCCTCA

mu.wsx.ecd                1645 GATACAGTACAGCATTGTGAGTGACCGAGTTAGCAAAGTTACCTTCTCCA
wsxfull.13.2.variant      1260 AAGCCAGTATGATGTTGTGAGTGATCATGTTAGCAAAGTTACTTTTTCA

mu.wsx.ecd                1695 ACCTGAAAGCCACCAGACCTCGAGGGAAGTTTACCTATGACGCAGTGTAC
wsxfull.13.2.variant      1310 ATCTGAATGAAACCAAACCTCGAGGAAAGTTTACCTATGATGCAGTGTAC

mu.wsx.ecd                1745 TGCTGCAATGAGCAGGCGTGCCATCACCGCTATGCTGAATTATACGTGAT
wsxfull.13.2.variant      1360 TGCTGCAATGAACATGAATGCCATCATCGCTATGCTGAATTATATGTGAT

mu.wsx.ecd                1795 CGATGTCAATATCAATATATCATGTGAAACTGACGGGTACTTAACTAAAA
wsxfull.13.2.variant      1410 TGATGTCAATATCAATATCTCATGTGAAACTGATGGGTACTTAACTAAAA

mu.wsx.ecd                1845 TGACTTGCAGATGGTCACCCAGCACAATCCAATCACTAGTGGGAAGCACT
wsxfull.13.2.variant      1460 TGACTTGCAGATGGTCAACCAGTACAATCCAGTCACTTGCGGAAAGCACT
```

FIG. 5F

```
mu.wsx.ecd              1895  GTGCAGCTGAGGTATCACAGGTGCAGCCTGTATTGTCCTGATAGTCCATC
wsxfull.13.2.variant    1510  TTGCAATTGAGGTATCATAGGAGCAGCCTTACTGTTCTGATATTCCATC

mu.wsx.ecd              1945  TATTCATCCTACGTCTGAGCCCAAA-ACTGCGTCTTACAGAGAGACGGCTT
wsxfull.13.2.variant    1560  TATTCATCCCATATCTGAGCCCAAAGATTGCTATTTGCAGAGTGATGGTT

mu.wsx.ecd              1994  TTTATGAATGTGTTTTCCAGCCAATCTTTCTATTATCTGGCTATACAATG
wsxfull.13.2.variant    1610  TTTATGAATGCATTTTCCAGCCAATCTTCCTATTATCTGGCTACACAATG

mu.wsx.ecd              2044  TGGATTCAGGATCAACCATTCTTTAGGTTCACTTGACTCGCCACCAACGTT
wsxfull.13.2.variant    1660  TGGATT-AGGATCAATCACTCTCTAGGTTCACTTGACTCTCCACCAACAT

mu.wsx.ecd              2094  GTGTCCTTCCTGACTCCGTAGTAAAACCACTACCTCCATCTAACGTAAAA
wsxfull.13.2.variant    1709  GTGTCCTTCCTGATTCTGTGGTGAAGCCACTGCCTCCATCCAGTGTGAAA

mu.wsx.ecd              2144  GCAGAGATTACTGTAAACACTGGATTATTGAAAGTATCTTGGGAAAAGCC
wsxfull.13.2.variant    1759  GCAGAAATTACTATAAACATGGATTATTGAAAATATCTTGGGAAAAGCC
```

FIG. 5G

```
mu.wsx.ecd              2194  AGTCTTTCCGGAGAATAACCTTCAATTCCAGATTCGATATGGCTTAAGTG
wsxfull.13.2.variant    1809  AGTCTTTCCAGAGAATAACCTTCAATTCCAGATTCGCTATGGTTAAGTG

mu.wsx.ecd              2244  GAAAAGAAATACAATGGAAGACACATGAGGTATTCGATGCAAAGTCAAAG
wsxfull.13.2.variant    1859  GAAAAGAAGTACAATGGAAGATGTATGAGGTTTATGATGCAAAATCAAAA

mu.wsx.ecd              2294  TCTGCCAGCCTGCTGGTCAGACCTCTGTGCAGTCTATGTGGTCAGGT
wsxfull.13.2.variant    1909  TCTGTCAGTCTCCCAGTTCCAGACTTGTGTGCAGTCTATGCTGTTCAGGT

mu.wsx.ecd              2344  TCGCTGCCGGCGGTTGGATGGACTAGGATATTGGAGTAATTGGAGCAGTC
wsxfull.13.2.variant    1959  GCGCTGTAAGAGGCTAGATGGACTGGGATATTGGAGTAATTGGAGCAATC

mu.wsx.ecd              2394  CAGCCTATACGCTTGTCATGGATGTAAAAGTTCCTATGAGAGGGCCTGAA
wsxfull.13.2.variant    2009  CAGCCTACACAGTTGTCATGGATATAAAAGTTCCTATGAGAGGACCTGAA

mu.wsx.ecd              2444  TTTTGGAGAAAAATGGATGGGGACGTTACTAAAAAGGAGAGAAATGTCAC
wsxfull.13.2.variant    2059  TTTTGGAGAATAATGATGGAGATACTATGAAAAAGGAGAAAAATGTCAC
```

FIG. 5H

```
mu.wsx.ecd              2494 CTTGCTTTGGAAGCCCCTGACGAAAAATGACTCACTGTGTAGTGTGAGGA
wsxfull.13.2.variant    2109 TTTACTTTGGAAGCCCCTGATGAAAAATGACTCATTGTGCAGTGTTCAGA

mu.wsx.ecd              2544 GGTACGTGGTGAAGCATCGTACTGCCCACAATGGGACGTGGTCAGAAGAT
wsxfull.13.2.variant    2159 GATATGTGATAACCATCATACTTCCTGCAATGGAACATGGTCAGAAGAT

mu.wsx.ecd              2594 GTGGGAAATCGGACCCAATCTCACTTTCCTGTGGACAGAACCAGCGCACAC
wsxfull.13.2.variant    2209 GTGGGAAATCACACGAAATTCACTTTCCTGTGGACAGAGCAAGCACATAC

mu.wsx.ecd              2644 TGTTACAGTTCTGGCTGTCAATTCCCTCGGCGCTTCCCTTGTGAATTTTA
wsxfull.13.2.variant    2259 TGTTACGGTTCTGGCCATCAATTCAATTGGTGCTTCTGTTGCAATTTTA

mu.wsx.ecd              2694 ACCTTACCTTCTCATGGCCCATGAGTAAAGTGAGTGCTGTGGAGTCACTC
wsxfull.13.2.variant    2309 ATTTAACCTTTTCATGGCCTATGAGCAAAGTAAATATCGTGCAGTCACTC

mu.wsx.ecd              2744 AGTGCTTATCCCCTGAGCAGCAGCTGTGTCATCCTTTCCTGGACACTGTC
wsxfull.13.2.variant    2359 AGTGCTTATCCTTTAAACAGCAGTTGTGTGATTGTTTCCTGGATACTATC
```

FIG. 5I

```
mu.wsx.ecd              2794 ACCTGATGATTATAGTCTGTTATATCTGGTTATTGAATGGAAGATCCTTA
wsxfull.13.2.variant    2409 ACCCAGTGATTACAAGCTAATGTATTTATTATTGAGTGGAAAATCTTA

mu.wsx.ecd              2844 ATGAAGATGATGGAATGAAGTGGCT
wsxfull.13.2.variant    2459 ATGAAGATGGTGAAATAAAATGGCTTAGAATCTCTTCATCTGTTAAGAAG

wsxfull.13.2.variant    2509 TATTATATCCATGATCATTTTATCCCCATTGAGAAGTACCAGTTCAGTCT

wsxfull.13.2.variant    2559 TTACCCAATATTTATGGAAGGAGTGGGAAAACCAAAGATAATTAATAGTT

wsxfull.13.2.variant    2609 TCACTCAAGATGATATTGAAAAACACCAGAGTGATGCAGGTTTATATGTA

wsxfull.13.2.variant    2659 ATTGTGCCAGTAATTATTTCCTCTTCCATCTTATTGCTTGGAACATTATT

wsxfull.13.2.variant    2709 AATATCACACCAAAGAATGAAAAAGCTATTTTGGGAAGATGTTCCGAACC

wsxfull.13.2.variant    2759 CCAAGAATTGTTCCTGGGCACAAGGACTTAATTTTCAGAAGCCAGAAACG
```

FIG. 5J

wsxfull.13.2.variant 2809 T T T G A G C A T C T T T T T A T C A A G C A T A C A G C A T C A G T G A C A T G T G G T C C T C T

wsxfull.13.2.variant 2859 T C T T T T G G A G C C T G A A A C A A T T T C A G A A G A T A T C A G T G T T G A T A C A T C A T

wsxfull.13.2.variant 2909 G G A A A A A T A A A G A T G A G A T G A T G C C A A C A A C T G T G G T C T C T C T A C T T T C A

wsxfull.13.2.variant 2959 A C A A C A G A T C T T G A A A A G G G T T C T G T T T G T A T T A G T G A C C A G T T C A A C A G

wsxfull.13.2.variant 3009 T G T T A A C T T C T C T G A G G C T G A G G G T A C T G A G G T A A C C T A T G A G G A C G A A A

wsxfull.13.2.variant 3059 G C C A G A G A C A A C C C T T T G T T A A A T A C G C C A C G C T G A T C A G C A A C T C T A A A

wsxfull.13.2.variant 3109 C C A A G T G A A A C T G G T G A A G A A C A A G G G C T T A T A A A T A G T T C A G T C A C C A A

wsxfull.13.2.variant 3159 G T G C T T C T C T A G C A A A A A T T C T C C G T T G A A G G A T T C T T T C T C T A A T A G C T

wsxfull.13.2.variant 3209 C A T G G G A G A T A G A G G C C C A G G C A T T T T T T A T A T T A T C A G A T C A G C A T C C C

FIG. 5K

wsxfull.13.2.variant 3259 A A C A T A A T T T C A C C A C A C C T C A C A T T C T C A G A A G G A T T G G A T G A A C T T T T

wsxfull.13.2.variant 3309 G A A A T T G G A G G G A A A T T T C C C T G A A G A A A A T A A T G A T A A A A A G T C T A T C T

wsxfull.13.2.variant 3359 A T T A T T T A G G G G T C A C C T C A A T C A A A A A G A G A G A G A G T G G T G T G C T T T T G

wsxfull.13.2.variant 3409 A C T G A C A A G T C A A G G G T A T C G T G C C C A T T C C C A G C C C C C T G T T T A T T C A C

wsxfull.13.2.variant 3459 G G A C A T C A G A G T T C T C C A G G A C A G T T G C T C A C A C T T T G T A G A A A A T A A T A

wsxfull.13.2.variant 3509 T C A A C T T A G G A A C T T C T A G T A A G A A G A C T T T T G C A T C T T A C A T G C C T C A A

wsxfull.13.2.variant 3559 T T C C A A A C T T G T T C T A C T C A G A C T C A T A A G A T C A T G G A A A A C A A G A T G T G

wsxfull.13.2.variant 3609 T G A C C T A A C T G T G T A A T T T C A C T G A A G A A A C C T T C A G A T T T G T G T T A T A A

wsxfull.13.2.variant 3659 T G G G T A A T A T A A A G T G T A A T A G A T T A T A G T T G T G G G T G G G A G A G A G A A A A

FIG. 5L

wsxfull.13.2.variant 3709 G A A A C C A G A G T C A A A T T T G A A A A T A A T T G T T C C A A A T G A A T G T T G T C T G T

wsxfull.13.2.variant 3759 T T G T T C T C T C T T A G T A A C A T A G A C A A A A A A T T T G A G A A A G C C T T C A T A A G

wsxfull.13.2.variant 3809 C C T A C C A A T G T A G A C A C G C T C T T C T A T T T T A T T C C C A A G C T C T A G T G G G A

wsxfull.13.2.variant 3859 A G G T C C C T T G T T T C C A G C T A G A A A T A A G C C C A A C A G A C A C C A T C T T T T G T

wsxfull.13.2.variant 3909 G A G A T G T A A T T G T T T T T T T C A G A G G G C G T G T T G T T T T A C C T C A A G T T T T T G

wsxfull.13.2.variant 3959 T T T T G T A C C A A C A C A C A C A C A C A C A C A C A C A T T C T T A A C A C A T G T C C T T G T G

wsxfull.13.2.variant 4009 T G T T T T G A G A G T A T A T T A T G T A T T T A T A T T T T G T G C T A T C A G A C T G T A G G

wsxfull.13.2.variant 4059 A T T T G A A G T A G G A C T T T C C T A A A T G T T T A A G A T A A A C A G A A T T C

FIG. 5M

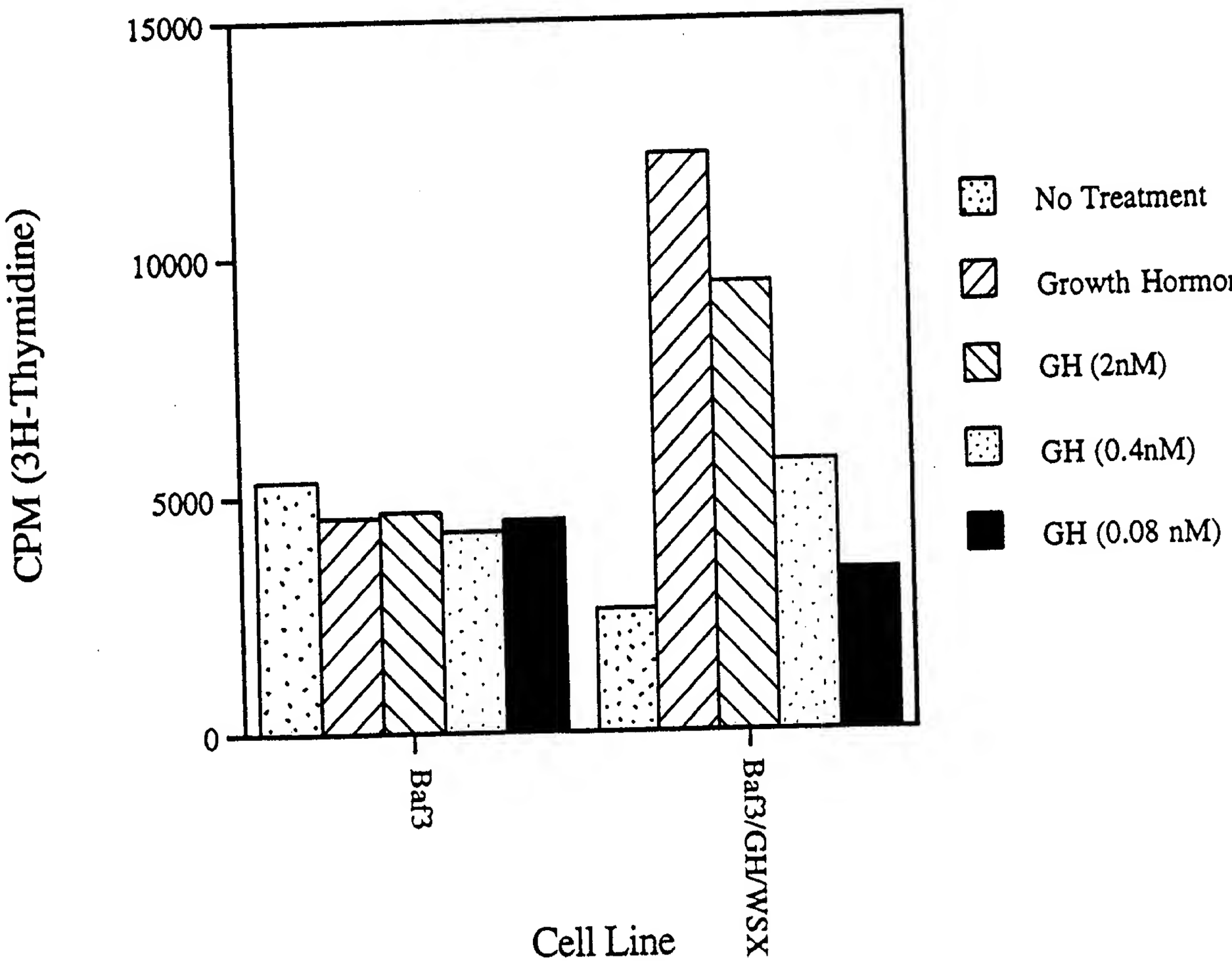

FIG. 6

<u>Murine</u>			
-213	Sense:	GGGTTAAGTTTCCCACCC	(SEQ ID NO:9)
	Antisense:	GGGTGGGAAACTTAACCC	(SEQ ID NO:10)
	Scrambled:	AGGATACAGTGGGATCCC	(SEQ ID NO:11)
-99	Sense:	GCCCGAGCACTCCTTTAA	(SEQ ID NO:12)
	Antisense:	TTAAAGGAGTGCTCCCGC	(SEQ ID NO:13)
	Scrambled:	GAGCGGCCCTGTTAGATA	(SEQ ID NO:14)
-20	Sense:	GTATACACCTCTGAAGAA	(SEQ ID NO:15)
	Antisense:	TTCTTCAGAGGTGTACAC	(SEQ ID NO:16)
	Scrambled:	ATGCGAGGCTACTTCTAT	(SEQ ID NO:17)
+84	Sense:	CTCTCCCTGGAAATTTAA	(SEQ ID NO:18)
	Antisense:	TTAAATTTCCAGGGAGAG	(SEQ ID NO:19)
	Scrambled:	ATTTGAAGGAGTTAAGCC	(SEQ ID NO:20)
+211	Sense:	AATTTAATTCAAGTGGTA	(SEQ ID NO:21)
	Antisense:	TACCAGTTGAATTAAATT	(SEQ ID NO:22)
	Scrambled:	GTATCACTTCATAATATA	(SEQ ID NO:23)
<u>Human</u>			
5L	Sense:	GATGGTCAGGGTGAACTG	(SEQ ID NO:24)
	Antisense:	CAGTTCACCCTGACCATC	(SEQ ID NO:25)
	Scrambled:	GAGGCGAATGTGCGGATT	(SEQ ID NO:26)
+85	Sense:	CTTAAATCTCCAAGGAGT	(SEQ ID NO:27)
	Antisense:	ACTCCTTGGAGATTTAAG	(SEQ ID NO:28)
	Scrambled:	AAGTCTTAAGCCAGACTT	(SEQ ID NO:29)
-47	Sense:	TCTAAGGCACATCCCAGC	(SEQ ID NO:30)
	Antisense:	GCTGGGATGTGCCTTAGA	(SEQ ID NO:31)
	Scrambled:	CGCAATGAATTGACCCCC	(SEQ ID NO:32)
-20	Sense:	TACTTCAGAGAAGTACAC	(SEQ ID NO:33)
	Antisense:	GTGTACTTCTCTGAAGTA	(SEQ ID NO:34)
	Scrambled:	GAATCACGGTAACTATCA	(SEQ ID NO:35)
+185	Sense:	CAGCTGTCTCATAATGTC	(SEQ ID NO:36)
	Antisense:	GACATTATGAGACAGCTG	(SEQ ID NO:37)
	Scrambled:	TTCGTCAAGCCATCTGAT	(SEQ ID NO:38)

FIG. 7

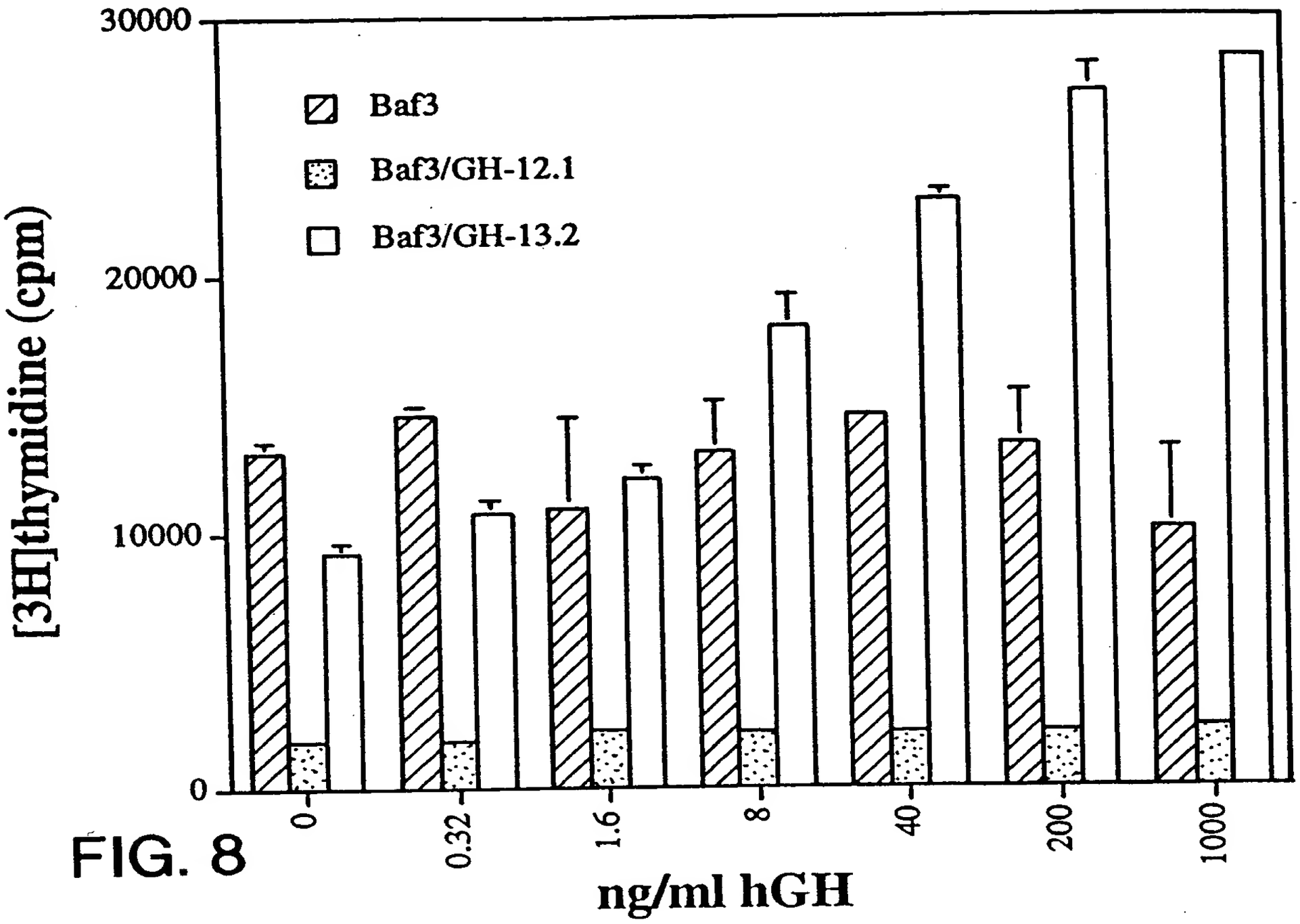

FIG. 8

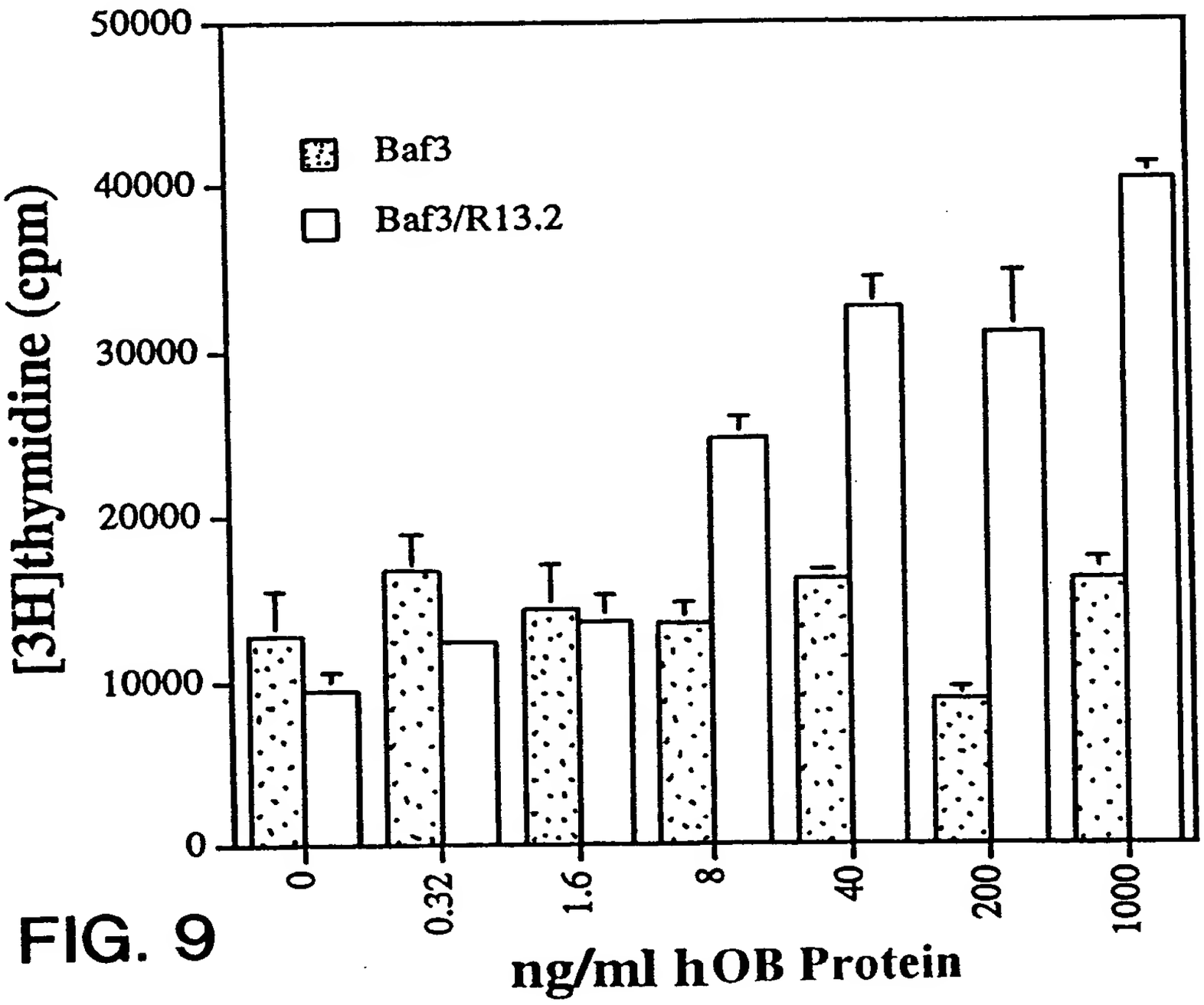

FIG. 9

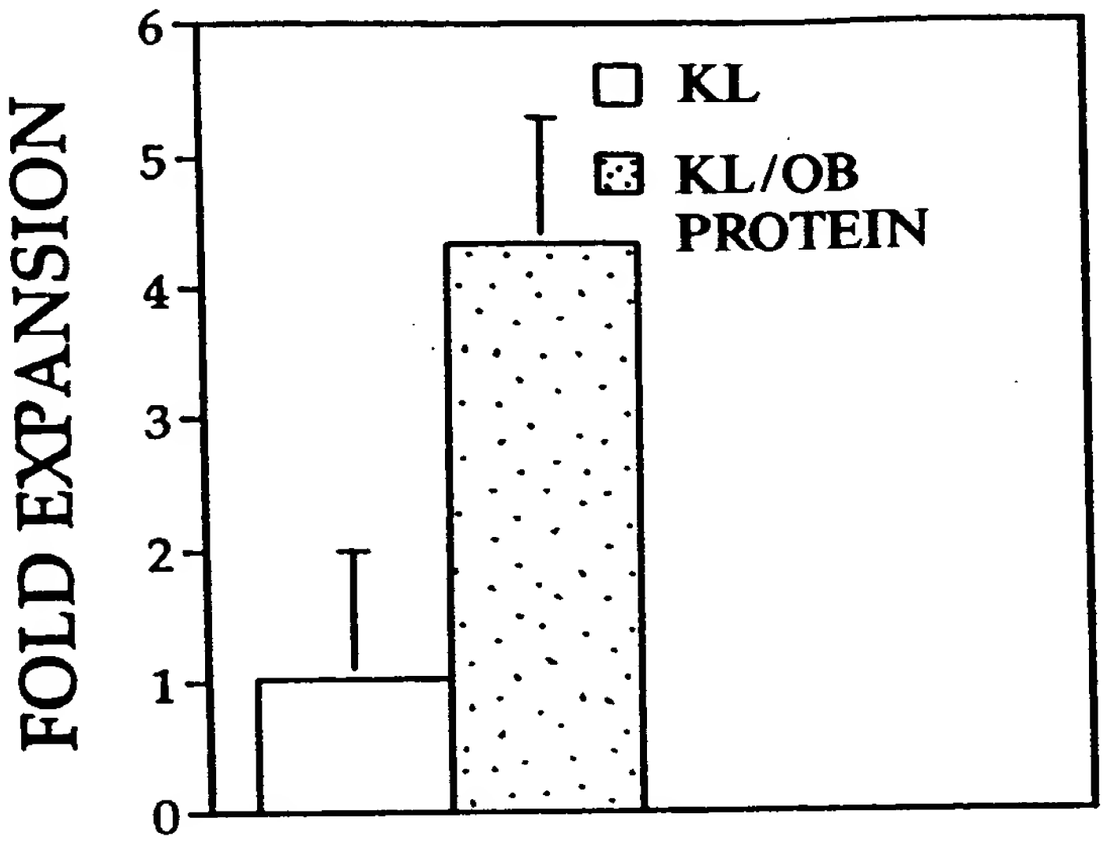

FIG. 10A

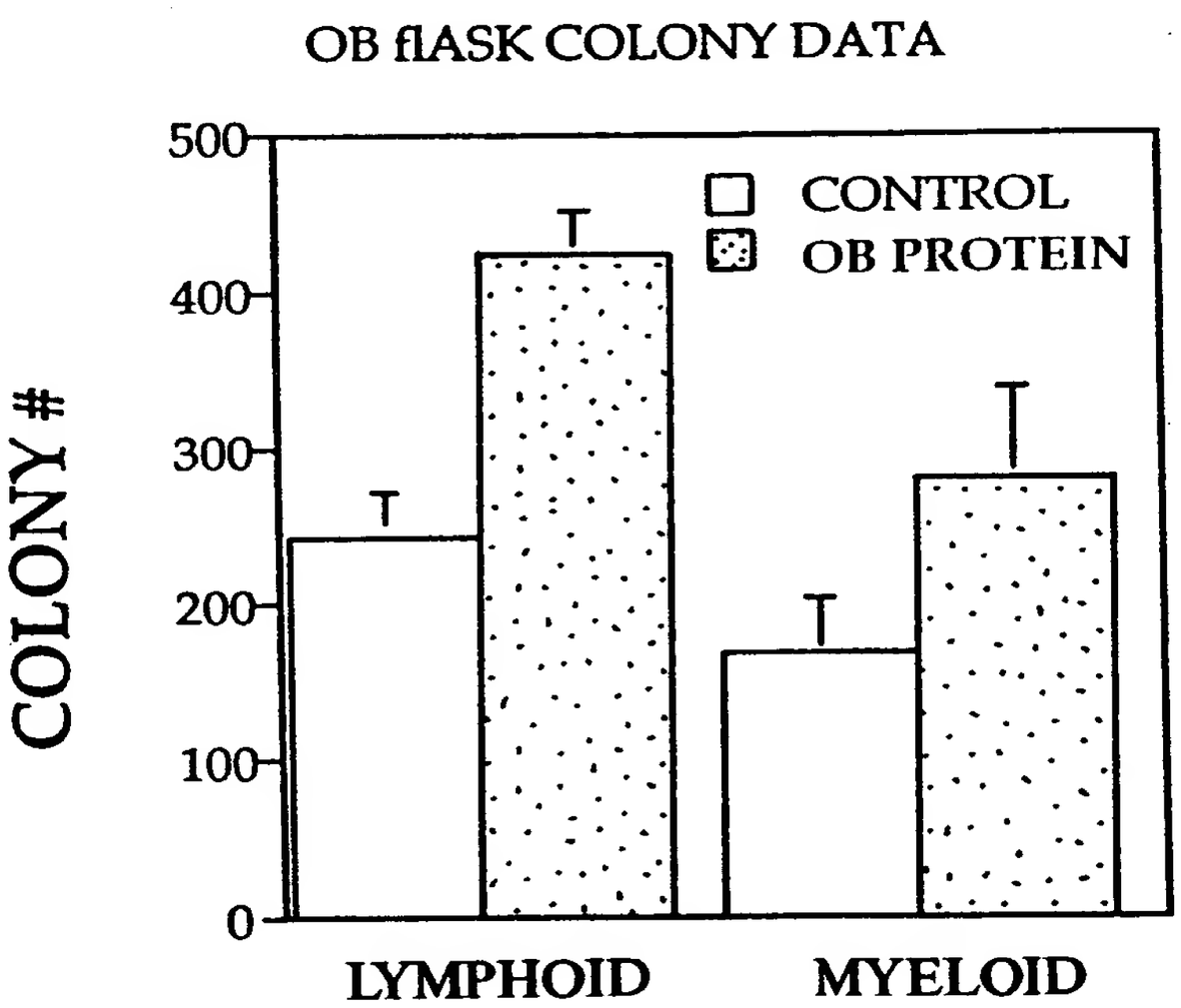

FIG. 10B

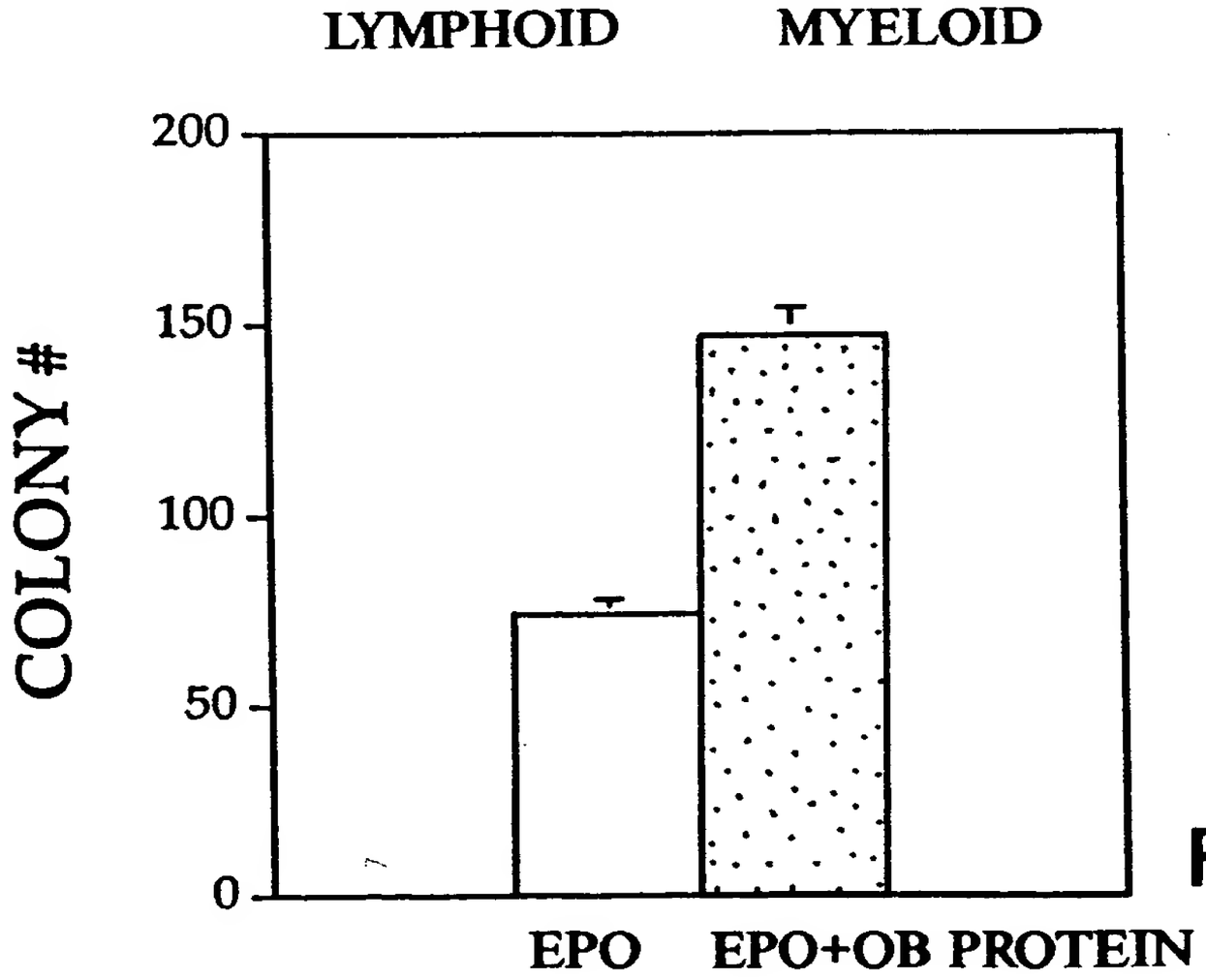

FIG. 10C

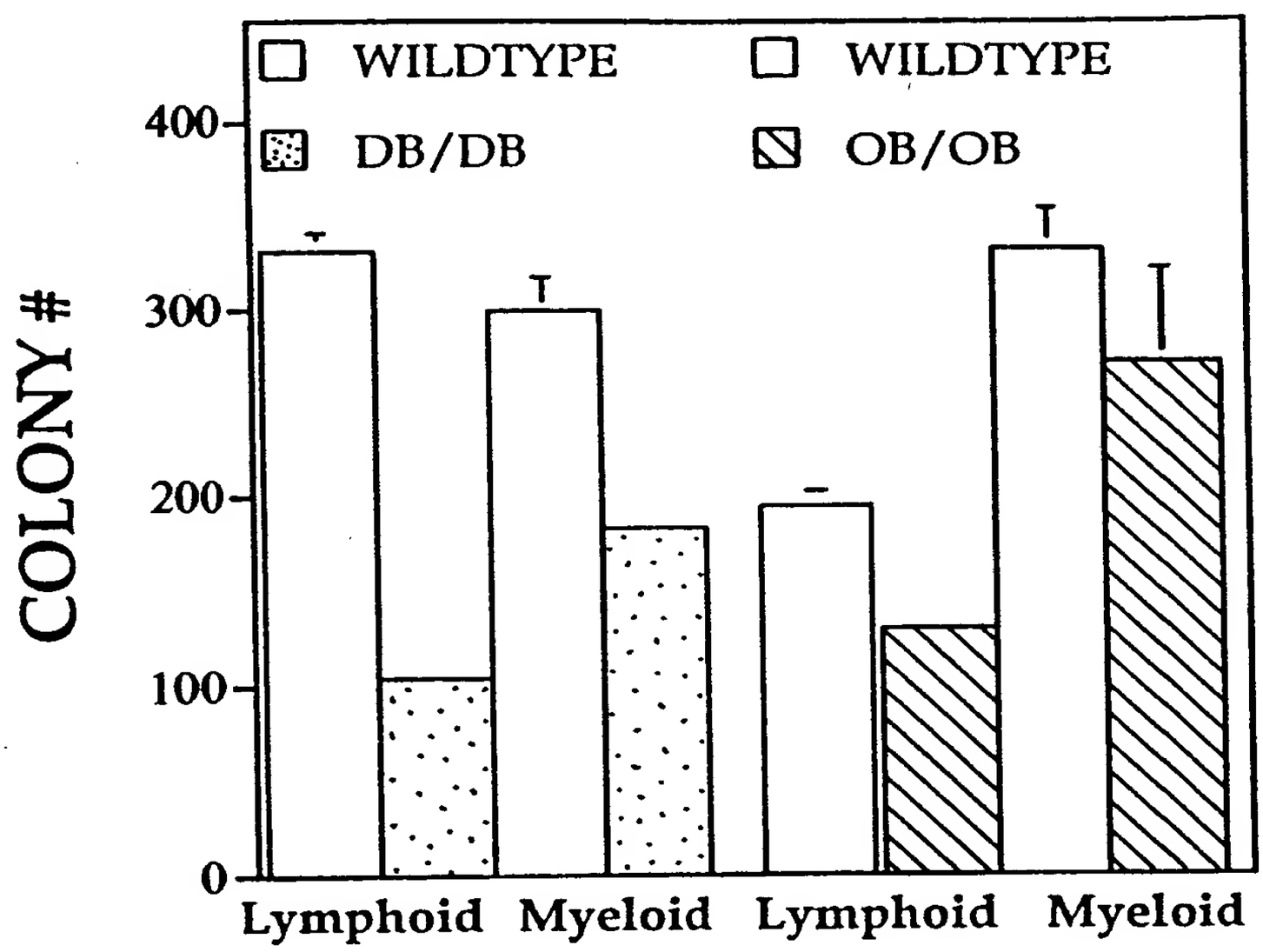

FIG. 11

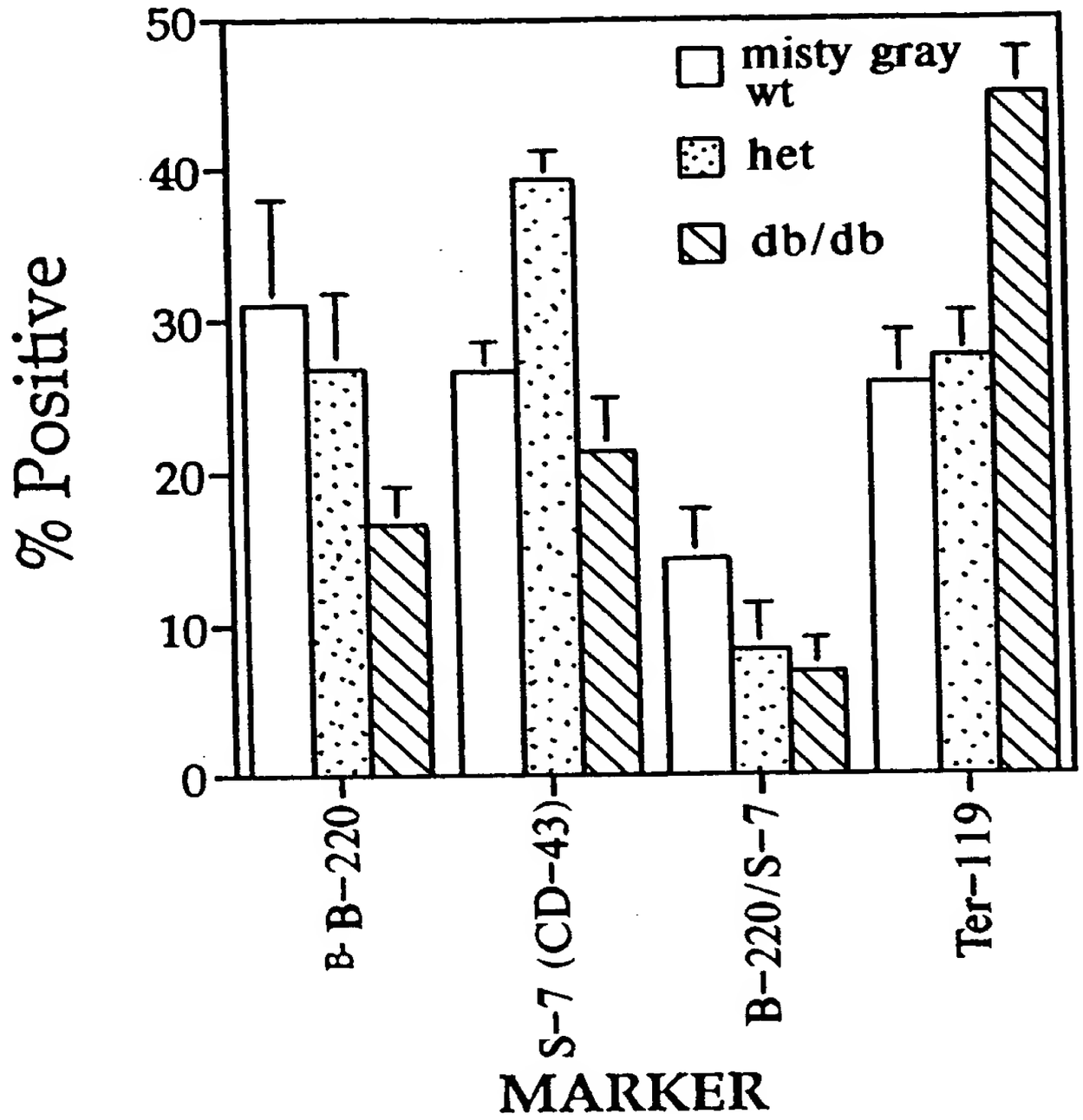

FIG. 12A

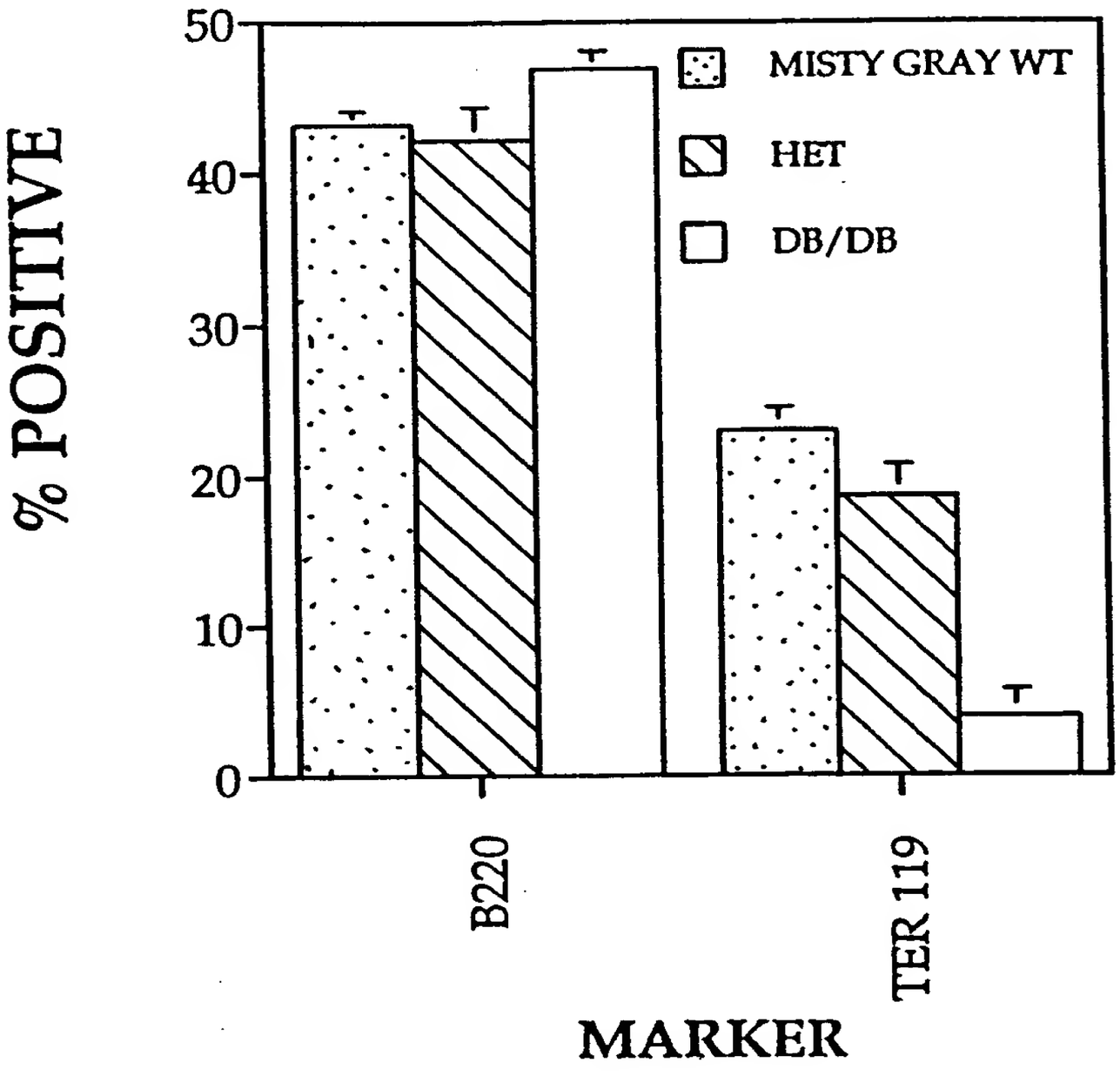

FIG. 12B

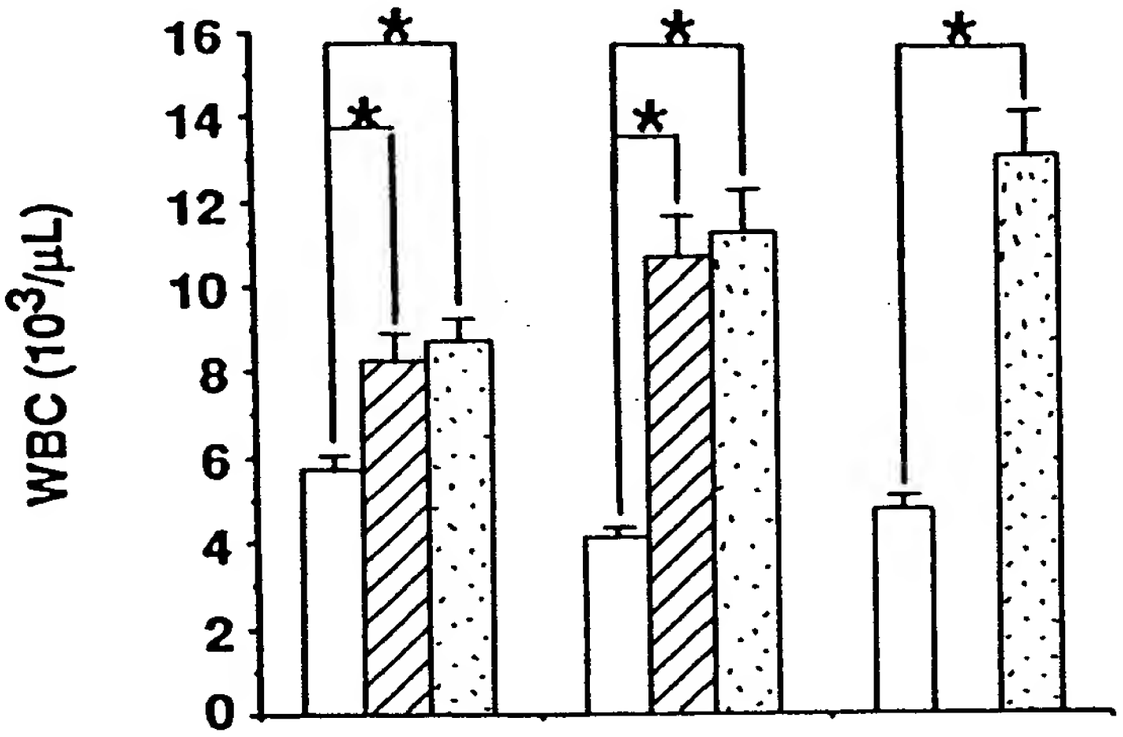

FIG. 13A

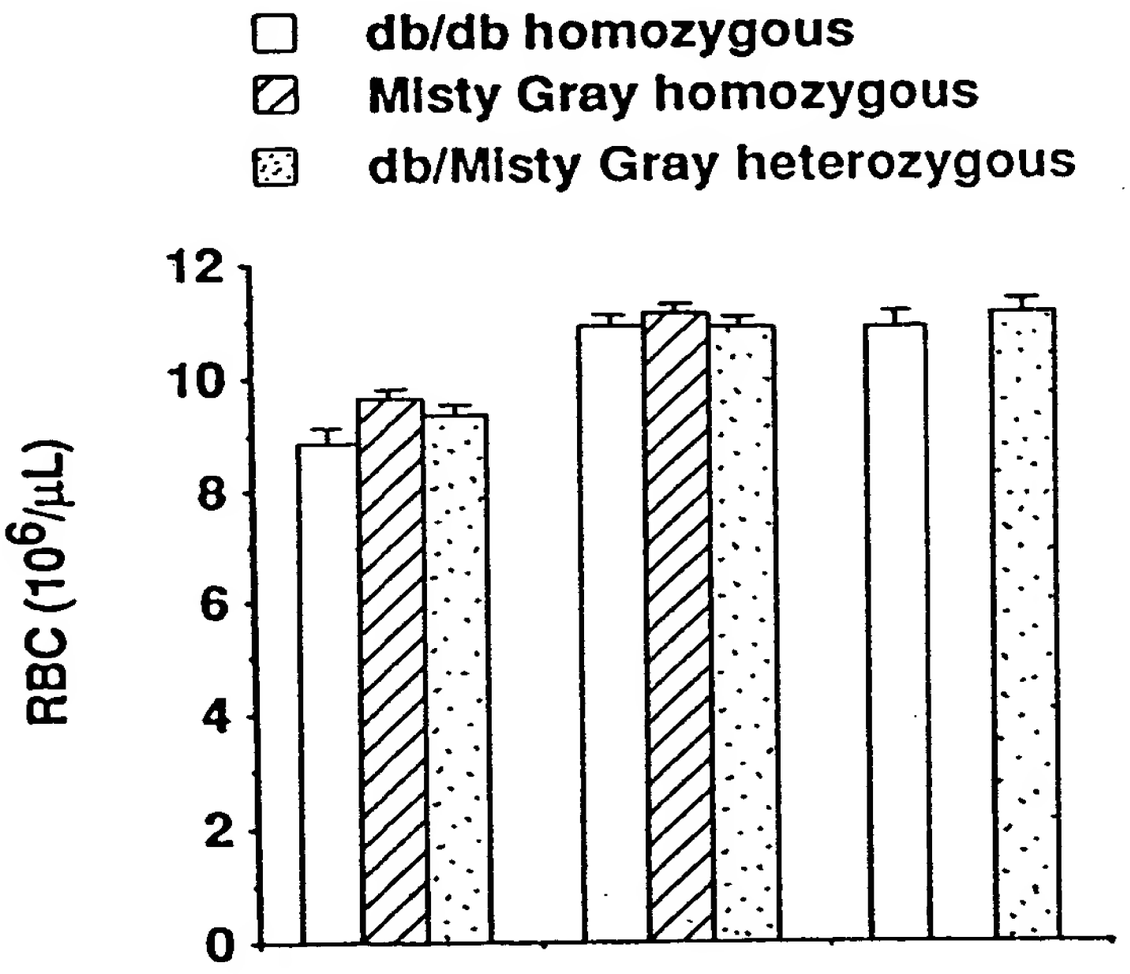

FIG. 13B

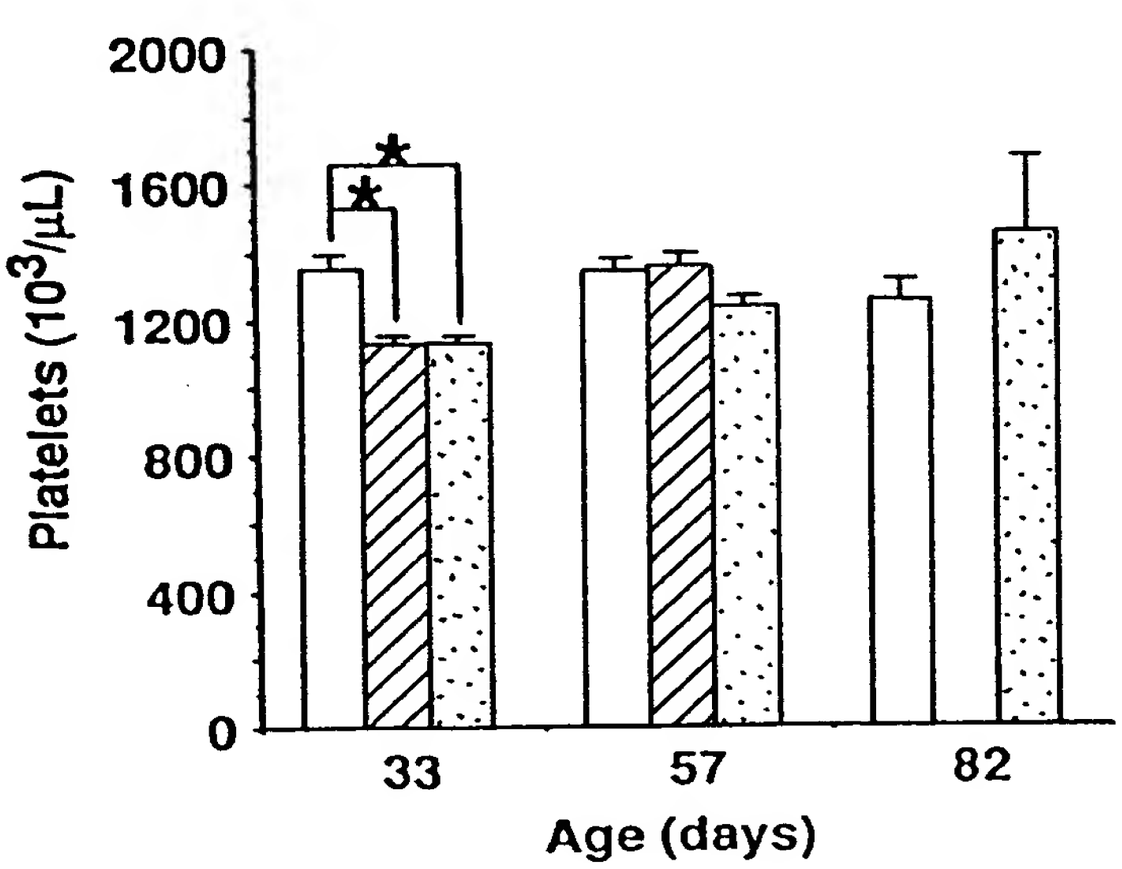

FIG. 13C

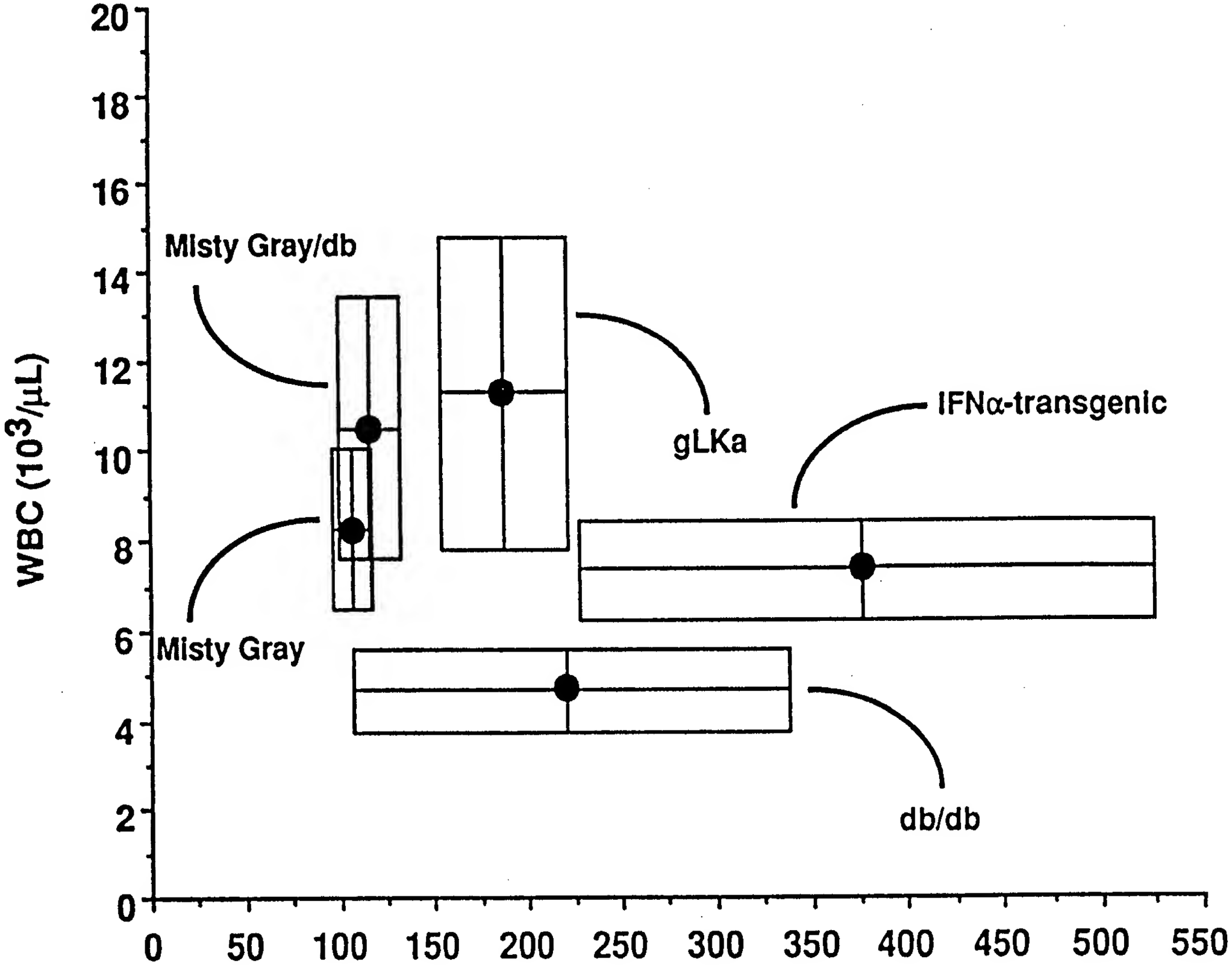

FIG. 14

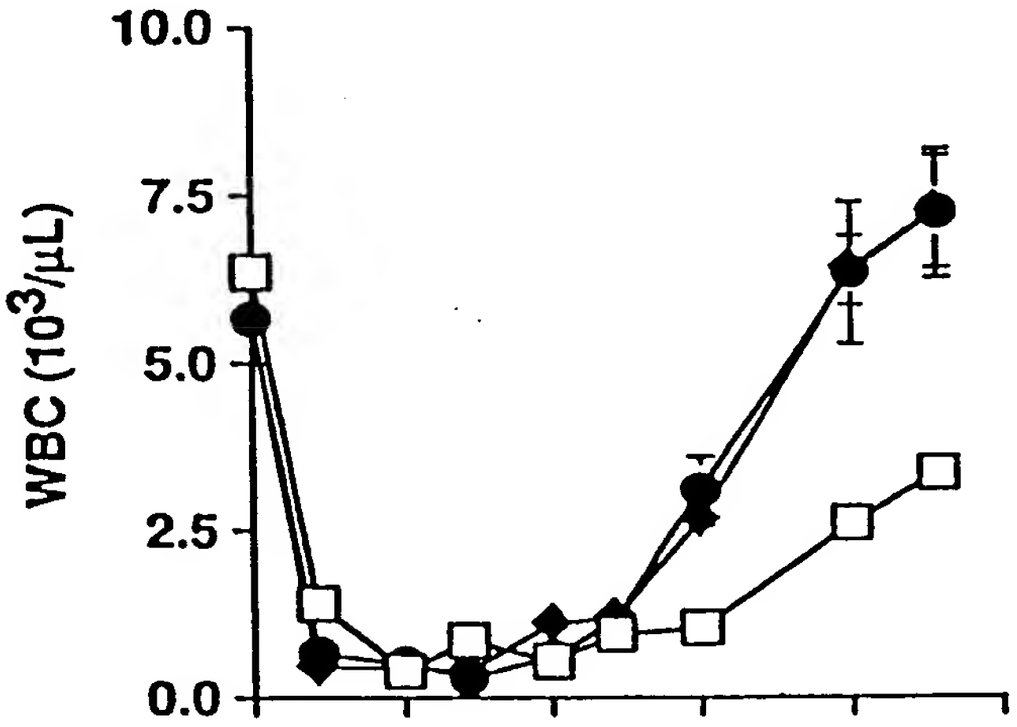

FIG. 15A

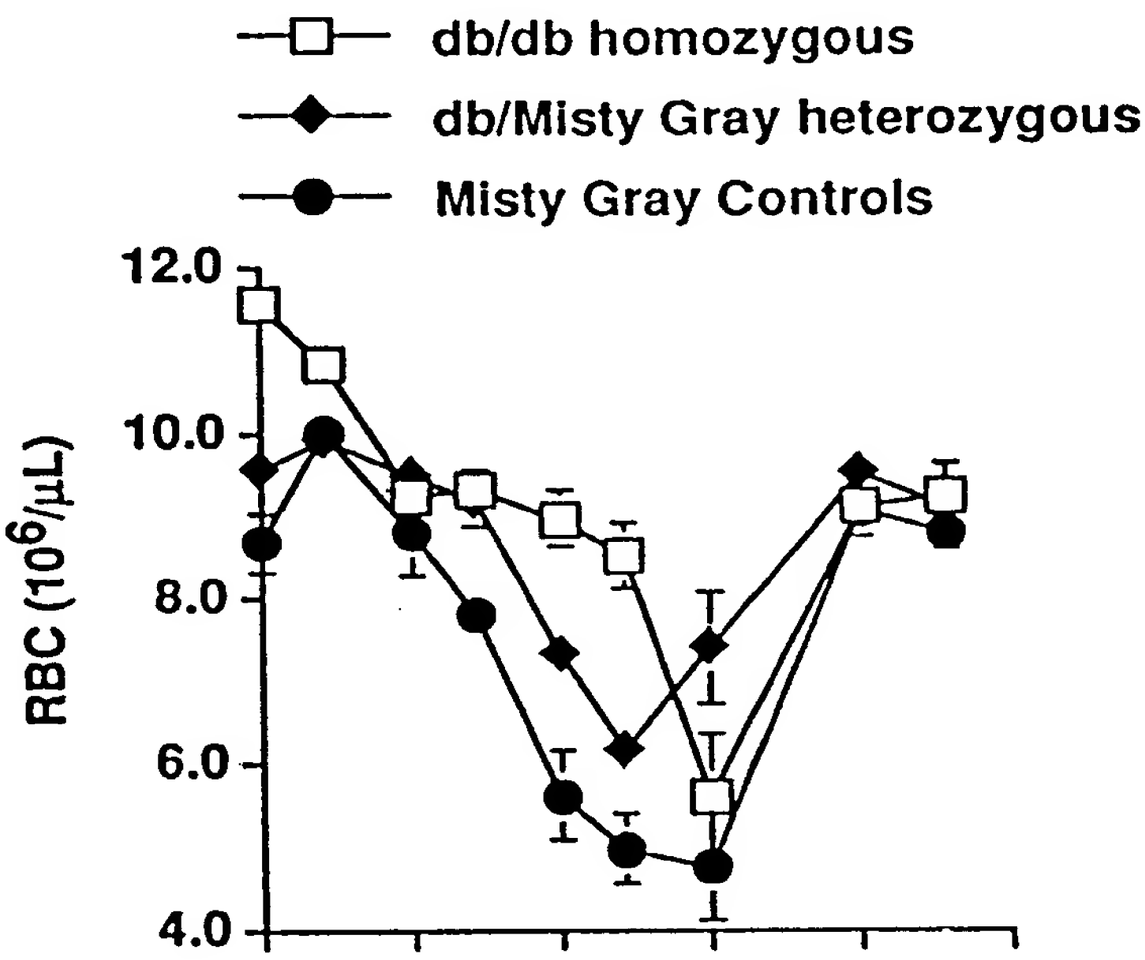

FIG. 15B

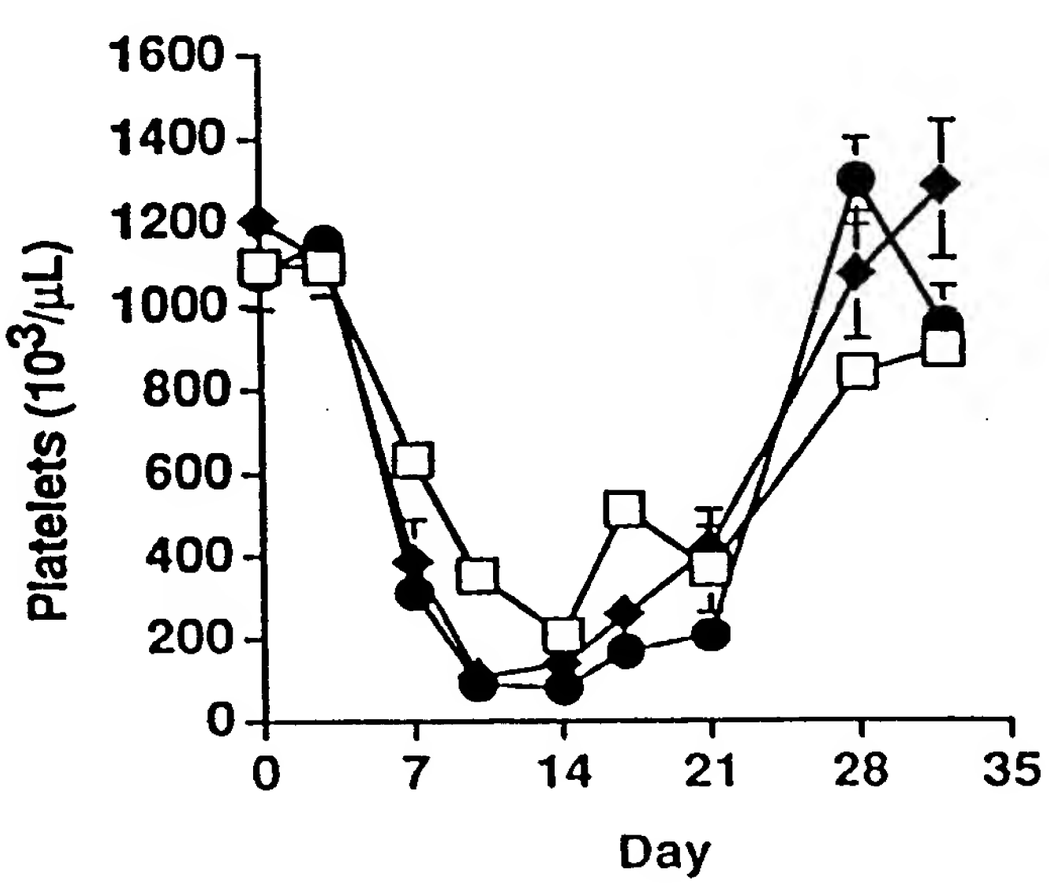

FIG. 15C

```
> sites: std
> length: 7127 (circular)

          aluI
         sstI
         sacI
         hgiJII
         hgiAI/aspHI
         ecl136II
         bsp1286                                                                    thaI
         bsiHKAI                                                                    fnuDII/mvnI
         bmyI                          rmaI     tru9I                               bstUI
         banII                         maeI     mseI                                bsh1236I
       taqI                           speI     aseI/asnI/vspI                 bslI  aciI maeIII
  1 TTCGAGCTCG CCCGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT TACGGGGTCA TTAGTTCATA GCCCATATAT GGAGTTCCGC GTTACATAAC
    AAGCTCGAGC GGGCTGTAAC TAATAACTGA TCAATAATTA TCATTAGTTA ATGCCCCAGT AATCAAGTAT CGGGTATATA CCTCAAGGCG CAATGTATTG

                    scrFI
                    mvaI
                    ecoRII
                    dsaV
                  aciI
                 bglI bstNI                          maeII
                sau96I                              hinlI/acyI
                haeIII/palI     aciI                ahaII/bsaHI
                asuI  apyI[dcm+]                aciI  aatII          maeII          maeIII
101 TTACGGTAAA TGGCCCGCCT GGCTGACCGC CCAACGACCC CCGCCCATTG ACGTCAATAA TGACGTATGT TCCCATAGTA ACGCCAATAG GGACTTTCCA
    AATGCCATTT ACCGGGCGGA CCGACTGGCG GGTTGCTGGG GGCGGGTAAC TGCAGTTATT ACTGCATACA AGGGTATCAT TGCGGTTATC CCTGAAAGGT

      maeII                                                                        maeII
     hinlI/acyI                                                                    hinlI/acyI
     ahaII/bsaHI                                  rsaI                  rsaI       ahaII/bsaHI
     aatII                              bglI      csp6I          ndeI   csp6I      aatII
201 TTGACGTCAA TGGGTGGAGT ATTTACGGTA AACTGCCCAC TTGGCAGTAC ATCAAGTGTA TCATATGCCA AGTACGCCCC CTATTGACGT CAATGACGGT
    AACTGCAGTT ACCCACCTCA TAAATGCCAT TTGACGGGTG AACCGTCATG TAGTTCACAT AGTATACGGT TCATGCGGGG GATAACTGCA GTTACTGCCA
```

FIG. 16A

```
                scrFI
                mvaI
                ecoRII                                                                                    nlaIII
             aciI                                                                                      styI
           bglI  dsaV                                                                                        ncoI
          sau96I bstNI              rsaI                                                    maeII           dsaI hphI  aciI
          haeIII/palI               csp6I                                           rsaI    snaBI           bsaJI   sfaNI
          asuI   apyI(dcm+)    bsrI  nlaIII                                        csp6I   bsaAI
301 AAATGGCCCG CCTGGCATTA TGCCCAGTAC ATGACCTTAT GGGACTTTCC TACTTGGCAG TACATCTACG TATTAGTCAT CGCTATTACC ATGGTGATGC
    TTTACCGGGC GGACCGTAAT ACGGGTCATG TACTGGAATA CCCTGAAAGG ATGAACCGTC ATGTAGATGC ATAATCAGTA GCGATAATGG TACCACTACG

                                                                                                maeII
                                                                                                hinII/acyI                nlaIV
                                                                                                ahaII/bsaHI               hgiCI
                   rsaI                                 pleI                                                          banI
                   csp6I                          aciI  hinfI                  bsmAI           aatII
401 GGTTTTGGCA GTACATCAAT GGGCGTGGAT AGCGGTTTGA CTCACGGGGA TTTCCAAGTC TCCACCCCAT TGACGTCAAT GGGAGTTTGT TTTGGCACCA
    CCAAAACCGT CATGTAGTTA CCCGCACCTA TCGCCAAACT GAGTGCCCCT AAAGGTTCAG AGGTGGGGTA ACTGCAGTTA CCCTCAAACA AAACCGTGGT

                                                                                                                          aluI
                                                                                                                          sstI
                                                                                                                          sacI
                                                                                                                          hgiJII
                                                                                                                          hgiAI/aspHI
                                                                                                                          ecl136II
                                                                                                                          bsp1286
                                                                                                                          bsiHKAI
                                                                                                    rsaI                  bmyI
                                  maeIII     aciI        hgaI          aciI          csp6I     mnlI                     banII
501 AAATCAACGG GACTTTCCAA AATGTCGTAA CAACTCCGCC CCATTGACGC AAATGGGCGG TAGGCGTGTA CGGTGGGAGG TCTATATAAG CAGAGCTCGT
    TTTAGTTGCC CTGAAAGGTT TTACAGCATT GTTGAGGCGG GGTAACTGCG TTTACCCGCC ATCCGCACAT GCCACCCTCC AGATATATTC GTCTCGAGCA
```

FIG. 16B

```
                                                                                                haeIII/palI
                                                                                               mcrI
                                                                                               eagI/xmaIII/eclXI
                                                                                               eaeI
                                                                                               cfrI
                                                                                              fnu4HI
                                                                                              aciI
                                                                                             thaI
                             esp3I                                                           fnuDII/mvnI
                        scrFI                                        sau96I           sacII/sstII
                        mvaI bsmAI                                         avaII            nspBII
                        ecoRII                                       asuI             kspI  scrFI
                        dsaV                                        nlaIV             dsaI  nciI
                        bstNI  hinlI/acyI                               scrFI          bglI bslI  mspI
                        apyI[dcm+]                                      nciI     sau3AI mnlI bstUI
                   sau3AI gsuI/bpmI                                     mspI     mboI/ndeII[dam-] hpaII
                   mboI/ndeII[dam-]                             mboII   hpaII    dpnI[dam+] bsaJI dsaV
                   dpnI[dam+]  hgaI  fokI                       bpuAI   dsaV     dpnII[dam-] bsh1236I
                   dpnII[dam-] ahaII/bsaHI             mnlI     bbsI    cauII    alwI[dam-] aciI  cauII
601 TTAGTGAACC GTCAGATCGC CTGGAGACGC CATCCACGCT GTTTTGACCT CCATAGAAGA CACCGGGACC GATCCAGCCT CCGCGGCCGG GAACGGTGCA
    AATCACTTGG CAGTCTAGCG GACCTCTGCG GTAGGTGCGA CAAAACTGGA GGTATCTTCT GTGGCCCTGG CTAGGTCGGA GGCGCCGGCC CTTGCCACGT
               ^Begin RNA

               tfiI                                                                              fnu4HI
           aciI                                                                                  aciI
          thaI hinfI                                                  bstXI                     thaI
          fnuDII/mvnI                         aciI                sau96I    styI                fnuDII/mvnI tru9I
          bstUI                     maeII rsaI          pleI scfI haeIII/palI                   bstUI      mseI
          bsh1236I        maeIII    csp6I   scfI hinfI     asuI      bsaJI                      bsh1236I  aseI/asnI/vspI
701 TTGGAACGCG GATTCCCCGT GCCAAGAGTG ACGTAAGTAC CGCCTATAGA GTCTATAGGC CCACCCCCTT GGCTTCGTTA GAACGCGGCT ACAATTAATA
    AACCTTGCGC CTAAGGGGCA CGGTTCTCAC TGCATTCATG GCGGATATCT CAGATATCCG GGTGGGGGAA CCGAAGCAAT CTTGCGCCGA TGTTAATTAT
                                                                                       ^sp6 promoter
```

FIG. 16C

```
                                                                                                                    sau96I
                                                                                                                    avaII
                                                                                                                    asuI
                                                                                                                  scrFI
                                                                                                                  mvaI
                                                                                                                  ecoRII
                                                                                                                  dsaV
                                                                                                                  bstNI
                                                                                                                  apyI[dcm+]
                                                maeIII
                                               hphI   scfI          fokI                                    bslI bsaJI
801 CATAACCTTA TGTATCATAC ACATACGATT TAGGTGACAC TATAGAATAA CATCCACTTT GCCTTTCTCT CCACAGGTGT CCACTCCCAG GTCCAACTGC
    GTATTGGAAT ACATAGTATG TGTATGCTAA ATCCACTGTG ATATCTTATT GTAGGTGAAA CGGAAAGAGA GGTGTCCACA GGTGAGGGTC CAGGTTGACG
                                                ^sp6 RNA start

                    ppu10I                    tfiI              sau96I
     mnlI        taqI nsiI/avaIII             hinfI             haeIII/palI                    bsp1286
     bsaJI      claI/bsp106     nlaIV      aciI                 asuI                      aluI bmyI fokI
901 ACCTCGGTTC TATCGATATG CATTGGGGAA CCCTGTGCGG ATTCTTGTGG CTTTGGCCCT ATCTTTTCTA TGTCCAAGCT GTGCCCATCC AAAAAGTCCA
    TGGAGCCAAG ATAGCTATAC GTAACCCCTT GGGACACGCC TAAGAACACC GAAACCGGGA TAGAAAAGAT ACAGGTTCGA CACGGGTAGG TTTTTCAGGT
  1                   Met HisTrpGlyT hrLeuCysGl yPheLeuTrp LeuTrpProT yrLeuPheTy rValGlnAla ValProIleG lnLysValGln
    ^cloning linker  ^human OB start

                                                sau3AI
                                                mboI/ndeII[dam-]
                                                dpnI[dam+]
                                              scrFI
                                              mvaI                                                          mspI
                                              ecoRII                                                      hpaII
                                              dsaV                                                        cfr10I
                                              bstNI                                                            bsaWI
                                              apyI[dcm+]                                                       ageI
                                            hphI   dpnII[dam-]                                 mnlI            hphI
                     mnlI        munI   maeIII alwI[dam-]                                  bsmAI             maeIII
1001 AGATGACACC AAAACCCTCA TCAAGACAAT TGTCACCAGG ATCAATGACA TTTCACACAC GCAGTCAGTC TCCTCCAAAC AGAAAGTCAC CGGTTTGGAC
     TCTACTGTGG TTTTGGGAGT AGTTCTGTTA ACAGTGGTCC TAGTTACTGT AAAGTGTGTG CGTCAGTCAG AGGAGGTTTG TCTTTCAGTG GCCAAACCTG
  29  AspAspThr LysThrLeuI leLysThrIl eValThrArg IleAsnAspI leSerHisTh rGlnSerVal SerSerLysG lnLysValTh rGlyLeuAsp
```

FIG. 16D

```
          hgiJII
          bsp1286
          bmyI
          banII
        scrFI
        mvaI nlaIV
        ecoRII
        dsaV
        bstNI
        bsaJI
        apyI[dcm+]      fokI              asuI       bsrI   accI     bstYI/xhoII               maeII
1101 TTCATTCCTG GGCTCCACCC CATCCTGACC TTATCCAAGA TGGACCAGAC ACTGGCAGTC TACCAACAGA TCCTCACCAG TATGCCTTCC AGAAACGTGA
     AAGTAAGGAC CCGAGGTGGG GTAGGACTGG AATAGGTTCT ACCTGGTCTG TGACCGTCAG ATGGTTGTCT AGGAGTGGTC ATACGGAAGG TCTTTGCACT
  62 PheIleProG lyLeuHisPr oIleLeuThr LeuSerLysM etAspGlnTh rLeuAlaVal TyrGlnGlnI leLeuThrSe rMetProSer ArgAsnValIle

(additional sites above 1101 row: pflMI, bslI, sau96I, avaII, alwNI; mnlI, sau3AI, bsrI, hphI, mboI/ndeII[dam-], dpnI[dam+], dpnII[dam-], alwI[dam-]; sau3AI, mboI/ndeII[dam-], dpnI[dam+], dpnII[dam-], alwI[dam-])

                                            pmlI                                             haeIII/palI
                                 sau3AI     eco72I                                            sau96I     styI
                                bstYI/xhoII                                                   asuI      gsuI/bpmI
           gsuI/bpmI    scrFI              bsaAI                                          scrFI    scrFI
           scrFI   ncII mboI/ndeII[dam-]                                                  mvaI           mvaI bsmAI
           mvaI         mspI dpnI[dam+]                                              ecoRII  ecoRII
           ecoRII hpaII     mboII maeII                                                   dsaV    dsaV
           dsaV   dsaV dpnII[dam-]                                fnu4HI             bstNI        bstNI    bsaJI
           bstNI     mnlI   alwI[dam-] bbrPI   haeIII/palI      bbvI                bsaJI  bstXI  apyI[dcm+]
      bstXI   apyI[dcm+]  cauII  mboII[dam-]   haeI      ddeI  aluI                    apyI[dcm+] haeI     bsaI
                                                                                     bsaJI    bsrI haeIII/palI
1201 TCCAAATATC CAACGACCTG GAGAACCTCC GGGATCTTCT TCACGTGCTG GCCTTCTCTA AGAGCTGCCA CTTGCCCTGG GCCAGTGGCC TGGAGACCTT
     AGGTTTATAG GTTGCTGGAC CTCTTGGAGG CCCTAGAAGA AGTGCACGAC CGGAAGAGAT TCTCGACGGT GAACGGGACC CGGTCACCGG ACCTCTGGAA
  96   GlnIleSe rAsnAspLeu GluAsnLeuA rgAspLeuLe uHisValLeu AlaPheSerL ysSerCysHi sLeuProTrp AlaSerGlyL euGluThrLeu

        scrFI           eco57I                                                                       sau96I
        mvaI          scrFI                                                                          avaII
        ecoRII mvaI    aluI                       ddeI      scfI                                   aluI
        dsaV   ecoRII                         bslI     pstI           scfI                   pvuII
        bstNI         dsaV                        sau96I        bsgI           pstI     nlaIII     nspBII
        bsaJI         bstNI hindIII     xcmI          haeIII/palI  fnu4HI     bsmAI      nspI       fnu4HI
        apyI[dcm+]    apyI[dcm+]        bstXI mnlI    asuI         bbvI       bsaI bsgI  nspHI      bbvI  asuI
1301 GGACAGCCTG GGGGGTGTCC TGGAAGCTTC AGGCTACTCC ACAGAGGTGG TGGCCCTGAG CAGGCTGCAG GGGTCTCTGC AGGACATGCT GTGGCAGCTG
     CCTGTCGGAC CCCCCACAGG ACCTTCGAAG TCCGATGAGG TGTCTCCACC ACCGGGACTC GTCCGACGTC CCCAGAGACG TCCTGTACGA CACCGTCGAC
 129  AspSerLeu GlyGlyValL euGluAlaSe rGlyTyrSer ThrGluValV alAlaLeuSe rArgLeuGln GlySerLeuG lnAspMetLe uTrpGlnLeu
```

FIG. 16E

```
                scrFI                                                                                                  eam1105I
                mvaI                                                                                                        sau96I
                ecoRII                                                                                        scrFI
                dsaV                                                                                          mvaI    avaII
                bstNI                                                                                               ecoRII
                bslI                                                                                                dsaV
                bsaJI           hphI                 nlaIII                                                         bstNI   asuI       mboII  mboII
         ddeI   apyI[dcm+]     maeIII               nspI          bsp1286                                          bsaJI  nlaIV      bpuAI  earI/ksp632I
        mnlI  bsaJI     aciI bstEII                 nspHI         bmyI alwNI                                       apyI[dcm+]        bbsI mnlI
1401 GACCTCAGCC CTGGGTGCGG GGTCACCGAC AAAACTCACA CATGCCCACC GTGCCCAGCA CCTGAACTCC TGGGGGGACC GTCAGTCTTC CTCTTCCCCC
     CTGGAGTCGG GACCCACGCC CCAGTGGCTG TTTTGAGTGT GTACGGGTGG CACGGGTCGT GGACTTGAGG ACCCCCCTGG CAGTCAGAAG GAGAAGGGGG
 162 AspLeuSerP roGlyCysGl yValThrAsp LysThrHisT hrCysProPr oCysProAla ProGluLeuL euGlyGlyPr oSerValPhe LeuPheProPro
                          ^insertion of a gly
                                  ^START OF HUMAN IgG1 CH2CH3

                                     sau96I
                                     nlaIV
                                   mspI
                                   hpaII
                                  scrFI
                                  nciI
                                  dsaV
                               sau3AI   avaII
                               mboI/ndeII[dam-]         nlaIII
                            nlaIII   cauII       mnlI   nspI                                                    drdI   mnlI                maeII
                           rcaI dpnI[dam+]     ddeI     nspHI                                                 mboII  ddeI                rsaI
          styI            bspHI[dam-]   asuI eco81I maeIII                                                  bpuAI eco81I               csp6I
          bsaJI      mnlI   dpnII[dam-]   bsu36I/mstII/sauI          maeII                                 bbsI  bsu36I/mstII/sauI bsrI bsaAI
1501 CAAAACCCAA GGACACCCTC ATGATCTCCC GGACCCCTGA GGTCACATGC GTGGTGGTGG ACGTGAGCCA CGAAGACCCT GAGGTCAAGT TCAACTGGTA
     GTTTTGGGTT CCTGTGGGAG TACTAGAGGG CCTGGGGACT CCAGTGTACG CACCACCACC TGCACTCGGT GCTTCTGGGA CTCCAGTTCA AGTTGACCAT
 196   LysProLy sAspThrLeu MetIleSerA rgThrProGl uValThrCys ValValValA spValSerHi sGluAspPro GluValLysP heAsnTrpTyr
```

FIG. 16F

```
                                                 aciI
                                                thaI
                                                fnuDII/mvnI
                                                bstUI
                                                bsh1236I
                                               sacII/sstII
                                               nspBII                                                                                scrFI
                                               kspI                                                                                  mvaI
                                               dsaI                                    rsaI                                          ecoRII
                                               bsaJI                                   csp6I                                         dsaV
                                               aciI                    rsaI       maeII                          hphI     ecoNI bstNI
                      mnlI                     fnu4HI  mnlI            csp6I      bsaAI                  hgaI mnlI      bslI  apyI[dcm+]
1601 CGTGGACGGC GTGGAGGTGC ATAATGCCAA GACAAAGCCG CGGGACGAGC AGTACAACAG CACGTACCGT GTGGTCAGCG TCCTCACCGT CCTGCACCAG
     GCACCTGCCG CACCTCCACG TATTACGGTT CTGTTTCGGC GCCCTCCTCG TCATGTTGTC GTGCATGGCA CACCAGTCGC AGGAGTGGCA GGACGTGGTC
 229  ValAspGly ValGluValH isAsnAlaLy sThrLysPro ArgGluGluG lnTyrAsnSe rThrTyrArg ValValSerV alLeuThrVa lLeuHisGln

                          rsaI          bsmAI                                                                    fnu4HI
     bsrI                 csp6I         bsaI            mnlI            taqI                                     bbvI  avaI
1701 GACTGGCTGA ATGGCAAGGA GTACAAGTGC AAGGTCTCCA ACAAAGCCCT CCCAGCCCCC ATCGAGAAAA CCATCTCCAA AGCCAAAGGG CAGCCCCGAG
     CTGACCGACT TACCGTTCCT CATGTTCACG TTCCAGAGGT TGTTTCGGGA GGGTCGGGGG TAGCTCTTTT GGTAGAGGTT TCGGTTTCCC GTCGGGGCTC
 262 AspTrpLeuA snGlyLysGl uTyrLysCys LysValSerA snLysAlaLe uProAlaPro IleGluLysT hrIleSerLy sAlaLysGly GlnProArgGlu

                                 scrFI
                                 nciI
                                 mspI
                                 hpaII
                                 dsaV
                                 cauII
                                xmaI/pspAI
                                smaI
                                scrFI                    scrFI          scrFI
                                nciI                     mvaI                mvaI
                                dsaV                     ecoRII         ecoRII
                                cauII                    dsaV           dsaV
                  rsaI          fokI                     bstNI               bstNI                                        dsaI
                  csp6I         bslI bsaJI mboII         apyI[dcm+]          apyI[dcm+]                                   bslI
                bsp1407I        bslI  avaI  earI/ksp632I  sexAI               bspMI                                        bsaJI
1801 AACCACAGGT GTACACCCTG CCCCCATCCC GGGAAGAGAT GACCAAGAAC CAGGTCAGCC TGACCTGCCT GGTCAAAGGC TTCTATCCCA GCGACATCGC
     TTGGTGTCCA CATGTGGGAC GGGGGTAGGG CCCTTCTCTA CTGGTTCTTG GTCCAGTCGG ACTGGACGGA CCAGTTTCCG AAGATAGGGT CGCTGTAGCG
 296   ProGlnVa lTyrThrLeu ProProSerA rgGluGluMe tThrLysAsn GlnValSerL euThrCysLe uValLysGly PheTyrProS erAspIleAla
```

FIG. 16G

```
                                  mspI
                                  hpaII                                                                                  dsaI
                              fnu4HI                                                 pleI                    mnlI          hphI
                              bbvI                                 mnlI              hinfI     nlaIV  mboII   scfI     aluI  bsaJI
1901 CGTGGAGTGG GAGAGCAATG GGCAGCCGGA GAACAACTAC AAGACCACGC CTCCCGTGCT GGACTCCGAC GGCTCCTTCT TCCTCTACAG CAAGCTCACC
     GCACCTCACC CTCTCGTTAC CCGTCGGCCT CTTGTTGATG TTCTGGTGCG GAGGGCACGA CCTGAGGCTG CCGAGGAAGA AGGAGATGTC GTTCGAGTGG
 329  ValGluTrp GluSerAsnG lyGlnProGl uAsnAsnTyr LysThrThrP roProValLe uAspSerAsp GlySerPheP heLeuTyrSe rLysLeuThr

                                                                                                                  scrFI
                                                                                                                  ncil
                                      mboII                                                                       mspI
                                     bpuAI                    nlaIII                                            hpaII
                                   maeII                     ppu10I                                 sapI         dsaV
                         fnu4HI    xmnI bbsI                 nsiI/avaIII                            mboII mnlI      bsmAI
                  bspMI  bbvI      asp700      nlaIII     sfaNI   mnlI                              earI/ksp632I bslI   cauII
2001 GTGGACAAGA GCAGGTGGCA GCAGGGGAAC GTCTTCTCAT GCTCCGTGAT GCATGAGGCT CTGCACAACC ACTACACGCA GAAGAGCCTC TCCCTGTCTC
     CACCTGTTCT CGTCCACCGT CGTCCCCTTG CAGAAGAGTA CGAGGCACTA CGTACTCCGA GACGTGTTGG TGATGTGCGT CTTCTCGGAG AGGGACAGAG
 362 ValAspLysS erArgTrpGl nGlnGlyAsn ValPheSerC ysSerValMe tHisGluAla LeuHisAsnH isTyrThrGl nLysSerLeu SerLeuSerPro

                                                                                                  sau96I
                                                                                                 nlaIII
                                    taqI                                                        fnu4HI   haeIII/palI
                                   salI                          taqI                           bglI styI
                                 pleI     scfI                pleI     scfI                   sfiI  ncoI
                               rmaI hincII/hindII             rmaI  salI  pstI                eaeI   dsaI
                          sau96I  hinfI    pstI              xbaI   hincII/hindII            cfrI   bsaJI                    aluI
                          haeIII/palI      bsgI    aluI maeI  accI  bsgI    aluI haeIII/palI                             fnu4HI
                          asuI maeI accI bspMI   hindIII   hinfI  bspMI    hindIII  aciI   asuI                          bbvI
2101 CGGGTAAATG AGTGCGACGG CCCTAGAGTC GACCTGCAGA AGCTTCTAGA GTCGACCTGC AGAAGCTTGG CCGCCATGGC CCAACTTGTT TATTGCAGCT
     GCCCATTTAC TCACGCTGCC GGGATCTCAG CTGGACGTCT TCGAAGATCT CAGCTGGACG TCTTCGAACC GGCGGTACCG GGTTGAACAA ATAACGTCGA
 396   GlyLys
                                                                                                  ^sv40 early poly A

                                                                                 rmaI
          maeIII                sfaNI   apoI                                  bsmI maeI
2201 TATAATGGTT ACAAATAAAG CAATAGCATC ACAAATTTCA CAAATAAAGC ATTTTTTTCA CTGCATTCTA GTTGTGGTTT GTCCAAACTC ATCAATGTAT
     ATATTACCAA TGTTTATTTC GTTATCGTAG TGTTTAAAGT GTTTATTTCG TAAAAAAAGT GACGTAAGAT CAACACCAAA CAGGTTTGAG TAGTTACATA
```

FIG. 16H

```
                    sau3AI
                    mboI/ndeII[dam-]
                    dpnI[dam+]
                    dpnII[dam-]
                   pvuI/bspCI
                   mcrI                                                                    rsaI
                  taqI[dam-]  tru9I                haeIII/palI                             csp6I
                 claI/bsp106[dam-]                 haeI                                    nlaIV
                sau3AI        mseI       fnu4HI styI                                      kpnI
                mboI/ndeII[dam-]         bbvI   ncoI                                      hgiCI
                dpnI[dam+] xmnI        hinPI    dsaI                                      banI
                dpnII[dam-]  aseI/asnI/vspI     bsaJI                                     asp718     mnlI
      nlaIII  alwI[dam-]  asp700      hhaI/cfoI nlaIII          mnlI      mnlI            acc65I  ddeI  aciI
2301  CTTATCATGT CTGGATCGAT CGGGAATTAA TTCGGCGCAG CACCATGGCC TGAAATAACC TCTGAAAGAG GAACTTGGTT AGGTACCTTC TGAGGCGGAA
      GAATAGTACA GACCTAGCTA GCCCTTAATT AAGCCGCGTC GTGGTACCGG ACTTTATTGG AGACTTTCTC CTTGAACCAA TCCATGGAAG ACTCCGCCTT
                          ^sv40 origin

                                                  nlaIV
                                                 scrFI                                  sfaNI                scrFI
                                                 mvaI                                       ppu10I                mvaI
                                                 ecoRII                                nsiI/avaIII           ecoRII
                                                 dsaV                                 nlaIII                 dsaV
                                                 bstNI                                     sphI                      bstNI
          aluI                                   apyI[dcm+]                                nspI                      apyI[dcm+]
         pvuII                                  bsaJI                                      nspHI                    sexAI
         nspBII
2401  AGAACCAGCT GTGGAATGTG TGTCAGTTAG GGTGTGGAAA GTCCCCAGGC TCCCCAGCAG GCAGAAGTAT GCAAAGCATG CATCTCAATT AGTCAGCAAC
      TCTTGGTCGA CACCTTACAC ACAGTCAATC CCACACCTTT CAGGGGTCCG AGGGGTCGTC CGTCTTCATA CGTTTCGTAC GTAGAGTTAA TCAGTCGTTG

                          nlaIV
                         scrFI
                         mvaI                                  ppu10I
                         ecoRII                              nsiI/avaIII
                         dsaV                               nlaIII
                         bstNI                                  sphI
                         apyI[dcm+]                             nspI sfaNI
                        bsaJI                                   nspHI                                  aciI          aciI fokI
2501  CAGGTGTGGA AAGTCCCCAG GCTCCCCAGC AGGCAGAAGT ATGCAAAGCA TGCATCTCAA TTAGTCAGCA ACCATAGTCC CGCCCCTAAC TCCGCCCATC
      GTCCACACCT TTCAGGGGTC CGAGGGGTCG TCCGTCTTCA TACGTTTCGT ACGTAGAGTT AATCAGTCGT TGGTATCAGG GCGGGGATTG AGGCGGGTAG
```

FIG. 16I

```
                                                                                          fnu4HI
                                                                                          bglI
                                                                                          sfiI
                                                  nlaIII                                  haeIII/palI
                                                  styI                                  mnlI     mnlI      ddeI
                                                  ncoI                                haeIII/palI bsaJI mnlI    aluI
                 bsrI                           bslI dsaI                           mnll bsaJI aciI    haeIII/palI
     aciI        aciI        aciI               aciI bsaJI
2601 CCGCCCCTAA CTCCGCCCAG TTCCGCCCAT TCTCCGCCCC ATGGCTGACT AATTTTTTTT ATTTATGCAG AGGCCGAGGC CGCCTCGGCC TCTGAGCTAT
     GGCGGGGATT GAGGCGGGTC AAGGCGGGTA AGAGGCGGGG TACCGACTGA TTAAAAAAAA TAAATACGTC TCCGGCTCCG GCGGAGCCGG AGACTCGATA

                                                                                                   scrFI
                                                                                                   mvaI
                                                                                                 sau96I
                                                                                                 nlaIV
                                                                                                 avaII   sau96I
                                                                                               thaI    ecoRII
                                                                                               fnuDII/mvnI
                                                                                               bstUI   dsaV
                                          styI                    sfuI                         bsh1236I   avaII
                                          bsaJI                   bstBI                        hinPI    bstNI
                                          blnI                    bsiCI                        hhaI/cfoI   asuI
                                        haeIII/pall               asuII                      fnu4HI asuI apyI[dcm+]
                                       stuI rmaI              tru9I                          aciI aciI  bsaJI
                 mnlI                  haeI maeI                                sfaNI
               mnlI                  mnlI avrII          aluI  mseI taqI
2701 TCCAGAAGTA GTGAGGAGGC TTTTTTGGAG GCCTAGGCTT TTGCAAAAAG CTGTTAATTC GAACACGCAG ATGCAGTCGG GGCGGCGCGG TCCCAGGTCC
     AGGTCTTCAT CACTCCTCCG AAAAAACCTC CGGATCCGAA AACGTTTTTC GACAATTAAG CTTGTGCGTC TACGTCAGCC CCGCCGCGCC AGGGTCCAGG
                                                       ^start pUC118
                                                              ^TK promoter

                                                                                                     sau3AI
                                                                                                     mboI/ndeII[dam-]
                                                                                                     dpnI[dam+]
                                                                                                 sau3AI
                                                                                                 mboI/ndeII[dam-]
                 thaI                                                                            dpnI[dam+]
                 fnuDII/mvnI                                                                    bstYI/xhoII
                 bstUI                                 fnu4HI                                   bglII  dpnII[dam-]
                mluI                                   bbvI                                      
                aflIII                                scfI
             hphI bsh1236I  mnlI                      pstI          tru9I   hincII/hindII aciI dpnII[dam-] bsmAI
          tru9I  hgaI     haeIII/palI                 bsgI      aciI mseI  hgaI         fnu4HI      bclI[dam-]
          mseI maeIII    haeI  taqI
2801 ACTTCGCATA TTAAGGTGAC GCGTGTGGCC TCGAACACCG AGCGACCCTG CAGCGACCCG CTTAACAGCG TCAACAGCGT GCCGCAGATC TGATCAAGAG
     TGAAGCGTAT AATTCCACTG CGCACACCGG AGCTTGTGGC TCGCTGGGAC GTCGCTGGGC GAATTGTCGC AGTTGTCGCA CGGCGTCTAG ACTAGTTCTC
                                                              tn5 neomycin phosphotransferase gene.   ^
```

FIG. 16J

```
                                                                     aciI
               sau3AI                                               fnu4HI
               mboI/ndeII[dam-]                                    haeIII/palI
               dpnI[dam+]                                         mcrI
               dpnII[dam-]                                        eagI/xmaIII/eclXI
            mnlI                                                  eaeI
          mamI[dam-]                                              cfrI                                 bsp1286
          bsaBI[dam-]                                            mspI                                  bmyI
         fokI  alwI[dam-]    nlaIII                  bspMI       hpaII          mnlI              bsrI
2901 ACAGGATGAG GATCGTTTCG CATGATTGAA CAAGATGGAT TGCACGCAGG TTCTCCGGCC GCTTGGGTGG AGAGGCTATT CGGCTATGAC TGGGCACAAC
     TGTCCTACTC CTAGCAAAGC GTACTAACTT GTTCTACCTA ACGTGCGTCC AAGAGGCCGG CGAACCCACC TCTCCGATAA GCCGATACTG ACCCGTGTTG

                                                        hinPI
                                                        hhaI/cfoI
                                                       nlaIV
                                                       narI scrFI                          bsp1286
                                                       kasI nciI                           bmyI
                                                       hinlI/acyI                         nlaIV
                                                       hgiCI mspI                         hgiCI
                                                       haeII hpaII                        banI
                           aciI                        banI dsaV                         mspI              scfI
           fnu4HI         fnu4HI      mspI     hinPI   ahaII/bsaHI                       hpaII             pstI
           bbvI     sfaNI    bslI     hpaII    hhaI/cfoI   cauII              drdI       bsaWI             bsgI
3001 AGACAATCGG CTGCTCTGAT GCCGCCGTGT TCCGGCTGTC AGCGCAGGGG CGCCCGGTTC TTTTTGTCAA GACCGACCTG TCCGGTGCCC TGAATGAACT
     TCTGTTAGCC GACGAGACTA CGGCGGCACA AGGCCGACAG TCGCGTCCCC GCGGGCCAAG AAAAACAGTT CTGGCTGGAC AGGCCACGGG ACTTACTTGA

                                                                 hgiAI/aspHI
               fnu4HI                                            bsp1286
              thaI                                            aluI
              fnuDII/mvnI                                    pvuII
              bstUI                                         fnu4HI
              bsh1236I        haeIII/palI                   bbvI  bsiHKAI
             hinPI           mscI/balI                     hinPI  bmyI
           fnu4HI            haeI                          hhaI/cfoI     maeII    eco57I
           bbvI aciI         eaeI                         mstI nspBII   tth111I/aspI                  fnu4HI
        mnlI  hhaI/cfoI      cfrI                         aviII/fspI   taqI     maeIII   aciI     bsrI bbvI
3101 GCAGGACGAG GCAGCGCGGC TATCGTGGCT GGCCACGACG GGCGTTCCTT GCGCAGCTGT GCTCGACGTT GTCACTGAAG CGGGAAGGGA CTGGCTGCTA
     CGTCCTGCTC CGTCGCGCCG ATAGCACCGA CCGGTGCTGC CCGCAAGGAA CGCGTCGACA CGAGCTGCAA CAGTGACTTC GCCCTTCCCT GACCGACGAT
```

FIG. 16K

```
                  scrFI
                  nciI       sau3AI                                                                    fnu4HI          sau3AI
                  mspI       mboI/ndeII[dam-]                                                        fnu4HI            mboI/ndeII[dam-]
                  hpaII      dpnI[dam+]                                                              aciI              dpnI[dam+]
                  dsaV       dpnII[dam-]                                                           fnu4HI              dpnII[dam-]
                  cauII     bstYI/xhoII                                                            aciI  bbvI          alwI[dam-]
                  bsaJI     alwI[dam-]       hphI                                    nlaIII  sfaNI
3201 TTGGGCGAAG TGCCGGGGCA GGATCTCCTG TCATCTCACC TTGCTCCTGC CGAGAAAGTA TCCATCATGG CTGATGCAAT GCGGCGGCTG CATACGCTTG
     AACCCGCTTC ACGGCCCCGT CCTAGAGGAC AGTAGAGTGG AACGAGGACG GCTCTTTCAT AGGTAGTACC GACTACGTTA CGCCGCCGAC GTATGCGAAC

                                                                                                    sau3AI
                                                                                                    mboI/ndeII[dam-]
                                                                                                fokI
                                                                                  sau3AI    dpnI[dam+]
                                                                                  mboI/ndeII[dam-]    sapI
                                                                                  dpnI[dam+]          mboII
                                                                                  dpnII[dam-]         earI/ksp632I
                                                                                  taqI[dam-]  dpnII[dam-]
                                                                  rsaI
                                                                  csp6I
                                                               bsaAI
                                                          hgiAI/aspHI
                                                          bsp1286              mspI
                                                          bsiHKAI              hpaII
                                                    taqI
       mspI
       hpaII bspMI      taqI                      sfaNI      bmyI maeII    fokI    cfr10I
3301 ATCCGGCTAC CTGCCCATTC GACCACCAAG CGAAACATCG CATCGAGCGA GCACGTACTC GGATGGAAGC CGGTCTTGTC GATCAGGATG ATCTGGACGA
     TAGGCCGATG GACGGGTAAG CTGGTGGTTC GCTTTGTAGC GTAGCTCGCT CGTGCATGAG CCTACCTTCG GCCAGAACAG CTAGTCCTAC TAGACCTGCT

                                                                  sphI
                                                                  nspI
                                                                  nspHI
                                                                hinPI
                                                                hhaI/cfoI
                                                              thaI                    sau3AI
                                                                                   mboI/ndeII[dam-]
                                                                                         dpnI[dam+]      styI
                                                                                   dpnII[dam-]     ncoI
                                                                                  bstYI/xhoII      dsaI
                                                                                        alwI[dam-]       bsaJI  sfaNI
                                                                                     mnlI          maeIII nlaIII
                   hinPI                                  scrFI  fnuDII/mvnI
                   thaI                                   mvaI         bstUI
                   fnuDII/mvnI                            ecoRII bsh1236I
                   bstUI                                  dsaV  hinPI nlaIII
                hgiJII                                    bstNI       hhaI/cfoI
                bsp1286                                   apyI[dcm+]  bssHII
                bmyI bsh1236I
       sfaNI    banII hhaI/cfoI
3401 AGAGCATCAG GGGCTCGCGC CAGCCGAACT GTTCGCCAGG CTCAAGGCGC GCATGCCCGA CGGCGAGGAT CTCGTCGTGA CCCATGGCGA TGCCTGCTTG
     TCTCGTAGTC CCCGAGCGCG GTCGGCTTGA CAAGCGGTCC GAGTTCCGCG CGTACGGGCT GCCGCTCCTA GAGCAGCACT GGGTACCGCT ACGGACGAAC
```

FIG. 16L

```
                                                                    mspI              bslI
                                                                    hpaII           sau96I
                                  aciI                            naeI              avaI
                                 fnu4HI                           cfr10I            asuI
                                haeIII/palI                      haeIII/palI       rsrII/cspI
                               eaeI           tfiI              eaeI              cpoI
              nlaIII           cfrI           hinfI  taqI       cfrI            aciI aciI
3501 CCGAATATCA TGGTGGAAAA TGGCCGCTTT TCTGGATTCA TCGACTGTGG CCGGCTGGGT GTGGCGGACC GCTATCAGGA CATAGCGTTG GCTACCCGTG
     GGCTTATAGT ACCACCTTTT ACCGGCGAAA AGACCTAAGT AGCTGACACC GGCCGACCCA CACCGCCTGG CGATAGTCCT GTATCGCAAC CGATGGGCAC

                sapI                                                                                  hinPI
                mboII      fnu4HI                                                   bsrBI             hhaI/cfoI
                earI/ksp632I                                                        aciI    tfiI  fnu4HI
              eco57I aluI  aciI              acrI   mnlI                          fnu4HI    hinfI bbvI sfaNI
3601 ATATTGCTGA AGAGCTTGGC GGCGAATGGG CTGACCGCTT CCTCGTGCTT TACGGTATCG CCGCTCCCGA TTCGCAGCGC ATCGCCTTCT ATCGCCTTCT
     TATAACGACT TCTCGAACCG CCGCTTACCC GACTGGCGAA GGAGCACGAA ATGCCATAGC GGCGAGGGCT AAGCGTCGCG TAGCGGAAGA TAGCGGAAGA

                                     taqI
                                    sfuI
                        aciI        bstBI                 hinlI/acyI                    tfiI       aciI
                    ddeI     pleI   bsiCI                 hgaI    bspMI                 hinfI     fnu4HI
                mboII  bsrBI  hinfI  asuII                 ahaII/bsaHI                 taqI      aciI
3701 TGACGAGTTC TTCTGAGCGG GACTCTGGGG TTCGAAATGA CCGACCAAGC GACGCCCAAC CTGCCATCAC GAGATTTCGA TTCCACCGCC GCCTTCTATG
     ACTGCTCAAG AAGACTCGCC CTGAGACCCC AAGCTTTACT GGCTGGTTCG CTGCGGGTTG GACGGTAGTG CTCTAAAGCT AAGGTGGCGG CGGAAGATAC
```

FIG. 16M

```
                                                                                                   scrFI
                                                                                                   nciI
                                                                                                   mspI
                                                                                                   hpaII
                                                                                                   dsaV
                                                                                                   cauII
                                                                                                   bslI
                               hinlI/acyI                      thaI  sau3AI                       xmaI/pspAI
                               hgaI                       bslI fnuDII/mvnI                        smaI
                               ahaII/bsaiII             gsuI/bpmI   bstYI/xhoII                   scrFI
                           scrFI                       mnlI   hinPI alwI[dam-]                    nciI
                           nciI                      sau3AI   hhaI/cfoI          gsuI/bpmI        dsaV
                           mspI     mspI             mboI/ndeII[dam-] mboI/ndeII[dam-]            cauII
                           hpaII     hpaII           dpnI[dam+] bstUI dpnI[dam+]                  bsaJI
                tfiI       dsaV     naeI             dpnII[dam-] aciI dpnII[dam-]                bsaJI
                hinfI      cauII    cfr10I fokI alwI[dam-] bshl236I      nlaIII        mboII   bslI avaI
3801 AAAGGTTGGG CTTCGGAATC GTTTTCCGGG ACGCCGGCTG GATGATCCTC CAGCGCGGGG ATCTCATGCT GGAGTTCTTC GCCCACCCCG GGAGATGGGG
     TTTCCAACCC GAAGCCTTAG CAAAAGGCCC TGCGGCCGAC CTACTAGGAG GTCGCGCCCC TAGAGTACGA CCTCAAGAAG CGGGTGGGGC CCTCTACCCC
                                                                   HSV1 tk terminator SmaI-PvuII  ^

                                                    hinPI
                                                    hhaI/cfoI
                                                   thaI
                                                   fnuDII/mvnI
                                        mspI       bstUI
                                         hpaII     bshl236I
     mnlI                    bsmAI      bsaWI  nlaIV aciI
3901 GAGGCTAACT GAAACACGGA AGGAGACAAT ACCGGAAGGA ACCCGCGCTA TGACGGCAAT AAAAAGACAG AATAAAACGC ACGGGTGTTG GGTCGTTTGT
     CTCCGATTGA CTTTGTGCCT TCCTCTGTTA TGGCCTTCCT TGGGCGCGAT ACTGCCGTTA TTTTTCTGTC TTATTTTGCG TGCCCACAAC CCAGCAAACA

                         scrFI
                         mvaI
                         ecoRII
                         dsaV
                         bstNI
                         bsaJI
                        bslI
        aciI        sau96I                                                                 thaI
       thaI         nlaIV                                                   haeIII/palI   fnuDII/mvnI
       fnuDII/mvnI     bsaJI                                      bslI       sau96I        bstUI
       bstUI        avaII                                      bsmAI        asuI          bshl236I
       bshl236I     asuI apyI[dcm+]        taqI                bsaI        nlaIV        aciI    mboII
4001 TCATAAACGC GGGGTTCGGT CCCAGGGCTG GCACTCTGTC GATACCCCAC CGAGACCCCA TTGGGGCCAA TACGCCCGCG TTTCTTCCTT TTCCCCACCC
     AGTATTTGCG CCCCAAGCCA GGGTCCCGAC CGTGAGACAG CTATGGGGTG GCTCTGGGGT AACCCCGGTT ATGCGGGCGC AAAGAAGGAA AAGGGGTGGG
```

FIG. 16N

```
                                    hgiJII                                                                  sau96I
                                    bsp1286                                                                 nlaIV
                                    bmyI                                                                    haeIII/palI
                                 scrFI                                                                  sau96I
                                 mvaI                                                                       nlaIV
                                 ecoRII                                                                 hgiJII                          bslI
                                 dsaV                                                                   eco0109I/draII              sau96I
                                 bstNI                                                                      bsp1286                         nlaIV
                                 bsaJI                                                                      bsp120I                         avaII styI
                                 apyI[dcm+]                                                                 bmyI                            asuI  ncoI
                                bsaJI                                                                       banII                          ppuMI  dsaI
                               sau96I     bslI                                                    xcmI asuI dsaI                  nlaIV  bsaJI
          bslI                 haeIII/palI  fnu4HI                fnu4HI                            dsaI apaI bsaJI                 eco0109I/draII
         bslI       hphI     asuI  banII  bbvI    maeII   aciI      aciI      bsaJI asuI                tth111I/aspI nlaIII
4101 CAACCCCCAA GTTCGGGTGA AGGCCCAGGG CTCGCAGCCA ACGTCGGGGC GGCAAGCCCG CCATAGCCAC GGGCCCCGTG GGTTAGGGAC GGGGTCCCCC
     GTTGGGGGTT CAAGCCCACT TCCGGGTCCC GAGCGTCGGT TGCAGCCCCG CCGTTCGGGC GGTATCGGTG CCCGGGGCAC CCAATCCCTG CCCCAGGGGG

                                                                                                                                nlaIII
                                                                                                                               styI
                                                                                sau96I                                         ncoI              haeIII/palI
                                                                                avaII                                          dsaI             haeI
                                                                                asuI     bsrI   ddeI                          bsaJI        fokI
4201 ATGGGGAATG GTTTATGGTT CGTGGGGGTT ATTCTTTTGG GCGTTGCGTG GGGTCAGGTC CACGACTGGA CTGAGCAGAC AGACCCATGG TTTTTGGATG
     TACCCCTTAC CAAATACCAA GCACCCCCAA TAAGAAAACC CGCAACGCAC CCCAGTCCAG GTGCTGACCT GACTCGTCTG TCTGGGTACC AAAAACCTAC

                                                                                                                                aciI
                                                                                                                               thaI
                                                                                                                               fnuDII/mvnI  hinPI
                                                                                                                               bstUI        hhaI/cfoI
      scrFI                                      hinlI/acyI                                                                 bsh1236I    nlaIV
      mvaI      sau96I                                ahaII/bsaHI                                                               hinPI        narI
      ecoRII               thaI          scrFI                                                                          thaI          kasI
      dsaV         rsaI    fnuDII/mvnI   nciI                                                                           fnuDII/mvnI   hinlI/acyI
      bstNI     avaII nlaIII     bstUI         mspI                                                                          bstUI         hgiCI
      bslI      asuI nspI       hinPI         hpaII                                                                         bsh1236I      haeII
      bsaJI nlaIII   nspHI  bsrI bsh1236I      dsaV hgaI      fnu4HI                                                        aciI          banI
      apyI[dcm+]   aciI  csp6I  hhaI/cfoI      cauII          bbvI                                                      bslI hhaI/cfoI    ahaII/bsaHI
4301 GCCTGGGCAT GGACCGCATG TACTGGCGCG ACACGAACAC CGGGCGTCTG TGGCTGCCAA ACACCCCCGA CCCCCAAAAA CCACCGCGCG GATTTCTGGC
     CGGACCCGTA CCTGGCGTAC ATGACCGCGC TGTGCTTGTG GCCCGCAGAC ACCGACGGTT TGTGGGGGCT GGGGGTTTTT GGTGGCGCGC CTAAAGACCG
```

FIG. 160

```
                                                                                                              aciI
                                                                                                             thaI
                                                                                                             fnuDII/mvnI
                                                                                                             bstUI
                                                                                                            sacII/sstII
                                                                                                  haeIII/palI bsh1236I
                                                                                                 mcrI        nspBII
                                                                                              dsaI           kspI
     mspI                                                                   hinPI             bsaJI          dsaI
     hpaII                                                          mnlI   hhaI/cfoI      hphI  eagI/xmaIII/eclXI
   aciI                                                        rsaI    haeII            maeIII  eaeI         bsaJI
   fnu4HI              bslI       sfaNI            mboII       csp6I   eco47III        bstEII  cfrI         aciI
4401 GCCGCCGGAC GAACTAAACC TGACTACGGC ATCTCTGCCC CTTCTTCGCT GGTACGAGGA GCGCTTTTGT TTTGTATTGG TCACCACGGC CGAGTTTCCG
     CGGCGGCCTG CTTGATTTGG ACTGATGCCG TAGAGACGGG GAAGAAGCGA CCATGCTCCT CGCGAAAACA AAACATAACC AGTGGTGCCG GCTCAAAGGC

       scrFI     nlaIV
       nciI      hgiCI
       dsaV  scrFI
       cauII mvaI
       bslI  ecoRII
      bslI   dsaV
      bsaJI  bstNI
    sau96I   bsaJI
    nlaIV  haeIII/palI
    avaII eaeI                                                                                      hinPI  mspI
    asuI  cfrI   bsp1286                                                                     thaI    hhaI/cfoI
   ppuMI mspI apyI(dcm+)                              mboII                                          hpaII
   nlaIV hpaII   bmyI                                 bpuAI                                        fnuDII/mvnI
   eco0109I/draII banI                                bbsI                                         bstUI  bsaWI
                                          nlaIII                     fnu4HI                        bsh1236I
                                                                     aciI    bcgI     nlaIII aciI   bslI        aluI
4501 CGGGACCCCG GCCAGGGCAC CTGTCCTACG AGTTGCATGA TAAAGAAGAC AGTCATAAGT GCGGCGACGA TAGTCATGCC CCGCGCCCAC CGGAAGGAGC
     GCCCTGGGGC CGGTCCCGTG GACAGGATGC TCAACGTACT ATTTCTTCTG TCAGTATTCA CGCCGCTGCT ATCAGTACGG GGCGCGGGTG GCCTTCCTCG
                 ^pBR322 sequence
```

FIG. 16P

```
                                                fnu4HI
                                                haeIII/palI                                                    aciI
                                                mcrI                                                          fnu4HI
                                                eagI/xmaIII/eclXI                hinPI                        thaI
                                                eaeI                             thaI                         fnuDII/mvnI
                                                notI                             fnuDII/mvnI                  bstUI
                                                fnu4HI                           bstUI scfI      hinPI      hinPI
                                                aciI                             bsh1236I        hhaI/cfoI  hhaI/cfoI
                                          mcrI  bsrBI aciI                     rsaI hhaI/cfoI  fnu4HI    tru9I aciI
      bsrI                           sfaNI    taqI cfrI  sfaNI                csp6I   bslI    aciI       mseI bsh1236I
4601  TGACTGGGTT GAAGGCTCTC AAGGGCATCG GTCGAGCGGC CGCATCAAAG CAACCATAGT ACGCGCCCTG TAGCGGCGCA TTAAGCGCGG CGGGTGTGGT
      ACTGACCCAA CTTCCGAGAG TTCCCGTAGC CAGCTCGCCG GCGTAGTTTC GTTGGTATCA TGCGCGGGAC ATCGCCGCGT AATTCGCGCC GCCCACACCA
                                                ^delta 3
                                                 ^M13 ori

            fnu4HI
            hinPI
            hhaI/cfoI                                  hinPI
            thaI                                       hhaI/cfoI
            fnuDII/mvnI                          rmaI                                                      mspI
            bstUI                                hinPI  haeII                                              hpaII
            bsh1236I      aciI                   hhaI/cfoI    bsrBI                                        naeI
       maeIII bbvI  maeIII                       haeII maeI    aciI            mboII                  maeII  cfr10I
4701  GGTTACGCGC AGCGTGACCG CTACACTTGC CAGCGCCCTA GCGCCCGCTC CTTTCGCTTT CTTCCCTTCC TTTCTCGCCA CGTTCGCCGG CTTTCCCCGT
      CCAATGCGCG TCGCACTGGC GATGTGAACG GTCGCGGGAT CGCGGGCGAG GAAAGCGAAA GAAGGGAAGG AAAGAGCGGT GCAAGCGGCC GAAAGGGGCA

                     nlaIV
                     hgiJII
                     bsp1286                                 nlaIV                                                 maeII    haeIII/palI
                     bmyI                                    hgiCI  taqI                                           draIII   sau96I
        aluI         banII          nlaIV                    banI mnlI                                hphI         bsaAI    asuI
4801  CAAGCTCTAA ATCGGGGGCT CCCTTTAGGG TTCCGATTTA GTGCTTTACG GCACCTCGAC CCCAAAAAAC TTGATTTGGG TGATGGTTCA CGTAGTGGGC
      GTTCGAGATT TAGCCCCCGA GGGAAATCCC AAGGCTAAAT CACGAAATGC CGTGGAGCTG GGGTTTTTTG AACTAAACCC ACTACCAAGT GCATCACCCG

                                      maeII pleI           tru9I     pleI                                    bslI
                                     drdI   hinfI  maeII   mseI      hinfI          bsrI                    bslI  avaI
4901  CATCGCCCTG ATAGACGGTT TTTCGCCCTT TGACGTTGGA GTCCACGTTC TTTAATAGTG GACTCTTGTT CCAAACTGGA ACAACACTCA ACCCTATCTC
      GTAGCGGGAC TATCTGCCAA AAAGCGGGAA ACTGCAACCT CAGGTGCAAG AAATTATCAC CTGAGAACAA GGTTTGACCT TGTTGTGAGT TGGGATAGAG
```

FIG. 16Q

```
                                                                                        thaI
                                                                                        fnuDII/mvnI
                                              tru9I                               tru9I apoI tru9I
                                              mseI              tru9I             mseI bstUI  mseI
                                  haeIII/palI           aluI    mseI       apoI   bsh1236I        sspI
5001 GGGCTATTCT TTTGATTTAT AAGGGATTTT GCCGATTTCG GCCTATTGGT TAAAAAATGA GCTGATTTAA CAAAAATTTA ACGCGAATTT TAACAAAATA
     CCCGATAAGA AAACTAAATA TTCCCTAAAA CGGCTAAAGC CGGATAACCA ATTTTTTACT CGACTAAATT GTTTTTAAAT TGCGCTTAAA ATTGTTTTAT

       maeII               mnlI                                                         maeII
       psp1406I            haeIII/palI                    nlaIII                        hinlI/acyI
     tru9I                 stuI                      tru9I rcaI                         ahaII/bsaHI
     mseI                  haeI                      mseI  bspHI              ddeI aatII
5101 TTAACGTTTA CAATTTTATG GTGCAGGCCT CGTGATACGC CTATTTTTAT AGGTTAATGT CATGATAATA ATGGTTTCTT AGACGTCAGG TGGCACTTTT
     AATTGCAAAT GTTAAAATAC CACGTCCGGA GCACTATGCG GATAAAAATA TCCAATTACA GTACTATTAT TACCAAAGAA TCTGCAGTCC ACCGTGAAAA
                           ^delta 2a

                nlaIV
                aciI
               thaI
               fnuDII/mvnI
               bstUI
               bsh1236I                                               rcaI
              hinPI                                                   bspHI
              hhaI/cfoI                                         bsrBI   bsmAI
                                                                aciI nlaIII                               sspI
5201 CGGGGAAATG TGCGCGGAAC CCCTATTTGT TTATTTTTCT AAATACATTC AAATATGTAT CCGCTCATGA GACAATAACC CTGATAAATG CTTCAATAAT
     GCCCCTTTAC ACGCGCCTTG GGGATAAACA AATAAAAAGA TTTATGTAAG TTTATACATA GGCGAGTACT CTGTTATTGG GACTATTTAC GAAGTTATTA

           mboII                                                 fnu4HI
           earI/ksp632I                                          aciI                  hphI
5301 ATTGAAAAAG GAAGAGTATG AGTATTCAAC ATTTCCGTGT CGCCCTTATT CCCTTTTTTG CGGCATTTTG CCTTCCTGTT TTTGCTCACC CAGAAACGCT
     TAACTTTTTC CTTCTCATAC TCATAAGTTG TAAAGGCACA GCGGGAATAA GGGAAAAAAC GCCGTAAAAC GGAAGGACAA AAACGAGTGG GTCTTTGCGA

                                 hgiAI/aspHI
                                 bsp1286                         sau3AI
                      sau3AI     bsiHKAI                         mboI/ndeII[dam-]  sau3AI
                      mboI/ndeII[dam-]                           dpnI[dam+]        mboI/ndeII[dam-]
                      dpnI[dam+] bmyI                            dpnII[dam-]       dpnI[dam+]
                      dpnII[dam-]                               bstYI/xhoII        dpnII[dam-]
                  eco57I         apaLI/snoI               bsrI         nspBII      alwI[dam-]
     hphI       sfaNI mboII[dam-] alw44I/snoI  maeIII taqI  alwI[dam-] aciI   bstYI/xhoII             mboII
5401 GGTGAAAGTA AAAGATGCTG AAGATCAGTT GGGTGCACGA GTGGGTTACA TCGAACTGGA TCTCAACAGC GGTAAGATCC TTGAGAGTTT TCGCCCCGAA
     CCACTTTCAT TTTCTACGAC TTCTAGTCAA CCCACGTGCT CACCCAATGT AGCTTGACCT AGAGTTGTCG CCATTCTAGG AACTCTCAAA AGCGGGGCTT
```

FIG. 16R

```
                                                                              scrFI
                                                                              nciI
                                                           aciI               mspI
                                                          thaI                hpaII
                                                          fnuDII/mvnI         dsaV
      maeII               hgiAI/aspHI                     bstUI               cauII
       psp1406I           bsp1286  tru9I                  bsh1236I          hinlI/acyI
      xmnI                bsiHKAI  mseI                  hinPI              hgaI                              aciI
      asp700              bmyI   ahaIII/draI             hhaI/cfoI          ahaII/bsaHI     bcgI    mcrI fnu4HI
5501 GAACGTTTTC CAATGATGAG CACTTTTAAA GTTCTGCTAT GTGGCGCGGT ATTATCCCGT GATGACGCCG GGCAAGAGCA ACTCGGTCGC CGCATACACT
     CTTGCAAAAG GTTACTACTC GTGAAAATTT CAAGACGATA CACCGCGCCA TAATAGGGCA CTACTGCGGC CCGTTCTCGT TGAGCCAGCG GCGTATGTGA

                                rsaI
                                csp6I  bsrI                                                          fnu4HI
        ddeI                   scaI hphI   maeIII       sfaNI      fokI  nlaIII                      bbvI       nlaIII
5601 ATTCTCAGAA TGACTTGGTT GAGTACTCAC CAGTCACAGA AAAGCATCTT ACGGATGGCA TGACAGTAAG AGAATTATGC AGTGCTGCCA TAACCATGAG
     TAAGAGTCTT ACTGAACCAA CTCATGAGTG GTCAGTGTCT TTTCGTAGAA TGCCTACCGT ACTGTCATTC TCTTAATACG TCACGACGGT ATTGGTACTC

                                                  sau96I
                                                  avaII                                           nlaIII
                                          sau3AI  asuI                                         sau3AI maeIII
                 haeIII/palI              mboI/ndeII[dam-]                                     mboI/ndeII[dam-]     sau3AI
                eaeI                      dpnI[dam+]                                           dpnI[dam+]           mboI/ndeII[dam-]
                cfrI                      dpnII[dam-]                                         alwI[dam-]            dpnI[dam+]
               fnu4HI                    pvuI/bspCI                                                                 dpnII[dam-]
               aciI                      mcrI  mnlI              aluI   aciI              nlaIII dpnII[dam-]
5701 TGATAACACT GCGGCCAACT TACTTCTGAC AACGATCGGA GGACCGAAGG AGCTAACCGC TTTTTTGCAC AACATGGGGG ATCATGTAAC TCGCCTTGAT
     ACTATTGTGA CGCCGGTTGA ATGAAGACTG TTGCTAGCCT CCTGGCTTCC TCGATTGGCG AAAAAACGTG TTGTACCCCC TAGTACATTG AGCGGAACTA

                                                                                              hinPI
                                                                                             mstI
              mspI                                                                           aviII/fspI    bsrI
              hpaII                                                                        maeII hhaI/cfoI tru9I
             bsaWI                                                           fnu4HI        psp1406I         mseI
          nlaIV    aluI                                    maeIII     sfaNI  bbvI
5801 CGTTGGGAAC CGGAGCTGAA TGAAGCCATA CCAAACGACG AGCGTGACAC CACGATGCCA GCAGCAATGG CAACAACGTT GCGCAAACTA TTAACTGGCG
     GCAACCCTTG GCCTCGACTT ACTTCGGTAT GGTTTGCTGC TCGCACTGTG GTGCTACGGT CGTCGTTACC GTTGTTGCAA CGCGTTTGAT AATTGACCGC
```

FIG. 16S

```
                      mspI
                      hpaII
                      scrFI
                 aluI nciI        tru9I     fokI                          sau96I          bglI
                                                                           sau96I
                rmaI  dsaV        mseI   bsrI       aciI                  avaII      hinPI asuI    mspI
                                                                                                   haeIII/palI
                maeI  cauII      aseI/asnI/vspI   mnlI                   asuI       hhaI/cfoI     hpaII
5901 AACTACTTAC TCTAGCTTCC CGGCAACAAT TAATAGACTG GATGGAGGCG GATAAAGTTG CAGGACCACT TCTGCGCTCG GCCCTTCCGG CTGGCTGGTT
     TTGATGAATG AGATCGAAGG GCCGTTGTTA ATTATCTGAC CTACCTCCGC CTATTTCAAC GTCCTGGTGA AGACGCGAGC CGGGAAGGCC GACCGACCAA

                                      aciI
                      mspI            thaI
                      hpaII           fnuDII/mvnI          sau96I
                     cfr10I           bstUI                asuI
                   nlaIV  hphI      bsmAI              fnu4HI   nlaIV
                                                                                                    eam1105I
                 gsuI/bpmI          bsaI bsh1236I      bbvI  bsrI haeIII/palI      mnlI
6001 TATTGCTGAT AAATCTGGAG CCGGTGAGCG TGGGTCTCGC GGTATCATTG CAGCACTGGG GCCAGATGGT AAGCCCTCCC GTATCGTAGT TATCTACACG
     ATAACGACTA TTTAGACCTC GGCCACTCGC ACCCAGAGCG CCATAGTAAC GTCGTGACCC CGGTCTACCA TTCGGGAGGG CATAGCATCA ATAGATGTGC

                                      ddeI
                                  sau3AI          nlaIV
                                  mboI/ndeII[dam-]   mnlI
                                  dpnI[dam+]      hgiCI          tru9I
          pleI                    dpnII[dam-]     banI           mseI       maeIII
          hinfI          fokI
6101 ACGGGGAGTC AGGCAACTAT GGATGAACGA AATAGACAGA TCGCTGAGAT AGGTGCCTCA CTGATTAAGC ATTGGTAACT GTCAGACCAA GTTTACTCAT
     TGCCCCTCAG TCCGTTGATA CCTACTTGCT TTATCTGTCT AGCGACTCTA TCCACGGAGT GACTAATTCG TAACCATTGA CAGTCTGGTT CAAATGAGTA

                                          hphI
                                      rmaI       sau3AI
                                  sau3AI         mboI/ndeII[dam-]
                                  mboI/ndeII[dam-]
                                  dpnI[dam+]     dpnI[dam+]
                              tru9I   dpnII[dam-]  dpnII[dam-]
                    tru9I           ahaIII/draI maeI      alwI[dam-]          nlaIII          maeII
                    mseI          tru9I       bstYI/xhoII  bstYI/xhoII       rcaI          tru9I
                   ahaIII/draI    mseI mseI   alwI[dam-] mboII[dam-]          bspHI         mseI
6201 ATATACTTTA GATTGATTTA AAACTTCATT TTTAATTTAA AAGGATCTAG GTGAAGATCC TTTTTGATAA TCTCATGACC AAAATCCCTT AACGTGAGTT
     TATATGAAAT CTAACTAAAT TTTGAAGTAA AAATTAAATT TTCCTAGATC CACTTCTAGG AAAAACTATT AGAGTACTGG TTTTAGGGAA TTGCACTCAA
```

FIG. 16T

```
                                                sau3AI
                                                mboI/ndeII[dam-]
                                                dpnI[dam+]  sau3AI          thaI
                                                dpnII[dam-] mboI/ndeII[dam-]
                                               bstYI/xhoII  dpnI[dam+]      fnuDII/mvnI
                                    sau3AI  alwI[dam-]    dpnII[dam-]      bstUI
                                    mboI/ndeII[dam-]      alwI[dam-]       bsh1236I
                                    dpnI[dam+] mboII[dam-]                 hinPI       fnu4HI
                                    dpnII[dam-]          bstYI/xhoII       hhaI/cfoI   bbvI
            ddeI hgaI
6301 TTCGTTCCAC TGAGCGTCAG ACCCCGTAGA AAAGATCAAA GGATCTTCTT GAGATCCTTT TTTTCTGCGC GTAATCTGCT GCTTGCAAAC AAAAAAACCA
     AAGCAAGGTG ACTCGCAGTC TGGGGCATCT TTTCTAGTTT CCTAGAAGAA CTCTAGGAAA AAAAGACGCG CATTAGACGA CGAACGTTTG TTTTTTGGT

                                sau3AI
                                mboI/ndeII[dam-]
                                dpnI[dam+]
                                dpnII[dam-]
                               alwI[dam-]
                                                                                                                rmaI
                                                                                                                maeI
                                                                                                                hinPI
                                                                                                                hhaI/cfoI
                                                                                                    bsrI
                                                                                                    maeIII   eco57I
                                                                                                                fnu4HI
                                                                                                                bbvI
                                                                                                    alwNI
                                                                                                    bsrI     fnu4HI
                                                                                                    maeIII   bbvI     bsrI
                                                                                                                mnlI
                                                                                                    scfI     aciI
                                                                                                    haeIII/palI
                                                                                                    haeI
                                                                                                    bslI
6501 TAGTGTAGCC GTAGTTAGGC CACCACTTCA AGAACTCTGT AGCACCGCCT ACATACCTCG CTCTGCTAAT CCTGTTACCA GTGGCTGCTG CCAGTGGCGA
     ATCACATCGG CATCAATCCG GTGGTGAAGT TCTTGAGACA TCGTGGCGGA TGTATGGAGC GAGACGATTA GGACAATGGT CACCGACGAC GGTCACCGCT

                                                                                                    hgiAI/aspHI
                                                                                                    bsp1286
                                                                                                    bsiHKAI
                                                                                                    bmyI
                                                                                                    apaLI/snoI
                                                                                                    alw44I/snoI    aluI
                                                                                                    hinPI aciI
                                                                                                    hhaI/cfoI
                                                                                                    fnu4HI
                                                                                                    bbvI
                                                                                                    nsp8II
                                                                                                    mcrI
                                                                                                    mspI
                                                                                                    hpaII
                                                                                                    bsaWI
                                                                                                    maeIII
                                    scrFI
                                    ncil
                                    mspI
                                    hpaII
                                    dsaV     pleI
                                    cauII    hinfI
6601 TAAGTCGTGT CTTACCGGGT TGGACTCAAG ACGATAGTTA CCGGATAAGG CGCAGCGGTC GGGCTGAACG GGGGGTTCGT GCACACAGCC CAGCTTGGAG
     ATTCAGCACA GAATGGCCCA ACCTGAGTTC TGCTATCAAT GGCCTATTCC GCGTCGCCAG CCCGACTTGC CCCCCAAGCA CGTGTGTCGG GTCGAACCTC
```

FIG. 16U

```
                                                                                                                    mspI
                                                                                                                    hpaII
                                                                                                                    bslI     fnu4HI
                                                                                                                    bsaWI    aciI
                                                                                                    aciI
                                                                                  hinPI
                                                                                  hhaI/cfoI
                                                                                  haeII
                                ddeI      scfI
6701 CGAACGACCT ACACCGAACT GAGATACCTA CAGCGTGAGC ATTGAGAAAG CGCCACGCTT CCCGAAGGGA GAAAGGCGGA CAGGTATCCG GTAAGCGGCA
     GCTTGCTGGA TGTGGCTTGA CTCTATGGAT GTCGCACTCG TAACTCTTTC GCGGTGCGAA GGGCTTCCCT CTTTCCGCCT GTCCATAGGC CATTCGCCGT

                                                  scrFI
                                                  mvaI          scrFI
                                                  ecoRII  mvaI
                                                  dsaV    ecoRII
                                                  bstNI           dsaV                                                taqI
                                hinPI  mnlI       bsaJI           bstNI
                                hhaI/cfoI  aluI   apyI[dcm+]      apyI[dcm+]                          mnlI  drdI    hgaI
6801 GGGTCGGAAC AGGAGAGCGC ACGAGGGAGC TTCCAGGGGG AAACGCCTGG TATCTTTATA GTCCTGTCGG GTTTCGCCAC CTCTGACTTG AGCGTCGATT
     CCCAGCCTTG TCCTCTCGCG TGCTCCCTCG AAGGTCCCCC TTTGCGGACC ATAGAAATAT CAGGACAGCC CAAAGCGGTG GAGACTGAAC TCGCAGCTAA

                                                                                                                    tru9I
                                                                                                                    mseI
                                                                                                  hinPI             aseI/asnI/vspI
                                                                                  bsrI    aciI    hhaI/cfoI
                                                  aluI
                                                  pvuII
                                nlaIV             nspBII
                      sfaNI     aciI
6901 TTTGTGATGC TCGTCAGGGG GGCGGAGCCT ATGGAAAAAC GCCAGCTGGC ACGACAGGTT TCCCGACTGG AAAGCGGGCA GTGAGCGCAA CGCAATTAAT
     AAACACTACG AGCAGTCCCC CCGCCTCGGA TACCTTTTTG CGGTCGACCG TGCTGTCCAA AGGGCTGACC TTTCGCCCGT CACTCGCGTT GCGTTAATTA
                                                  ^deltal.PVU

                                    scrFI
                                    mvaI
                                    ecoRII
                                    dsaV
                                nlaIV bstNI                                                         aciI
                      mnlI      hgiCI apyI[dcm+]                mspI                              bsrBI
                   maeIII       banI bsaJI                      hpaII
7001 GTGAGTTACC TCACTCATTA GGCACCCCAG GCTTTACACT TTATGCTTCC GGCTCGTATG TTGTGTGGAA TTGTGAGCGG ATAACAATTT CACACAGGAA
     CACTCAATGG AGTGAGTAAT CCGTGGGGTC CGAAATGTGA AATACGAAGG CCGAGCATAC AACACACCTT AACACTCGCC TATTGTTAAA GTGTGTCCTT

                                    tru9I
                                    mseI
                                    aseI/asnI/vspI
                                xmnI
                   aluI  nlaIII asp700
7101 ACAGCTATGA CCATGATTAC GAATTAA
     TGTCGATACT GGTACTAATG CTTAATT
```

FIG. 16V

>length: 7127

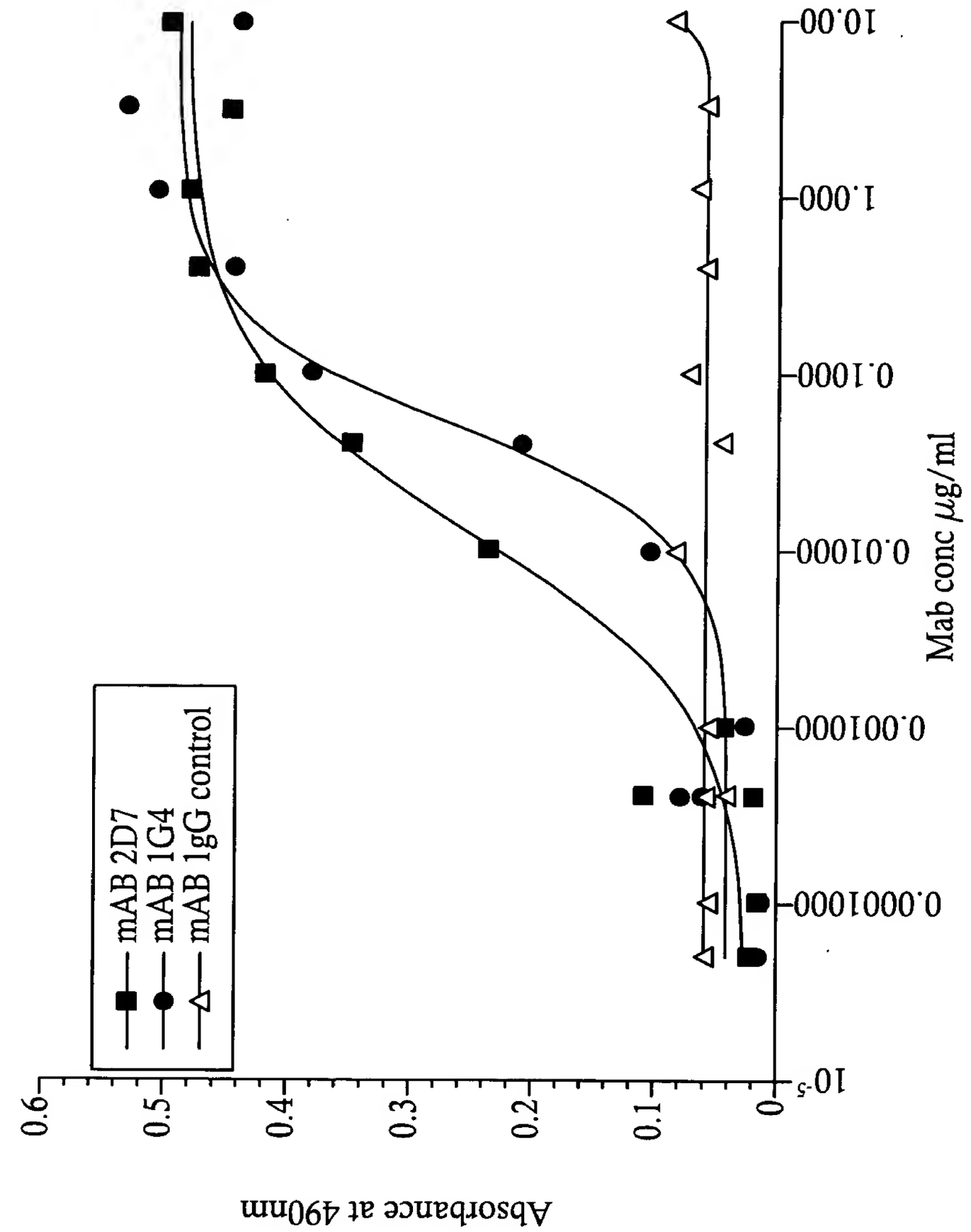

FIG. 17

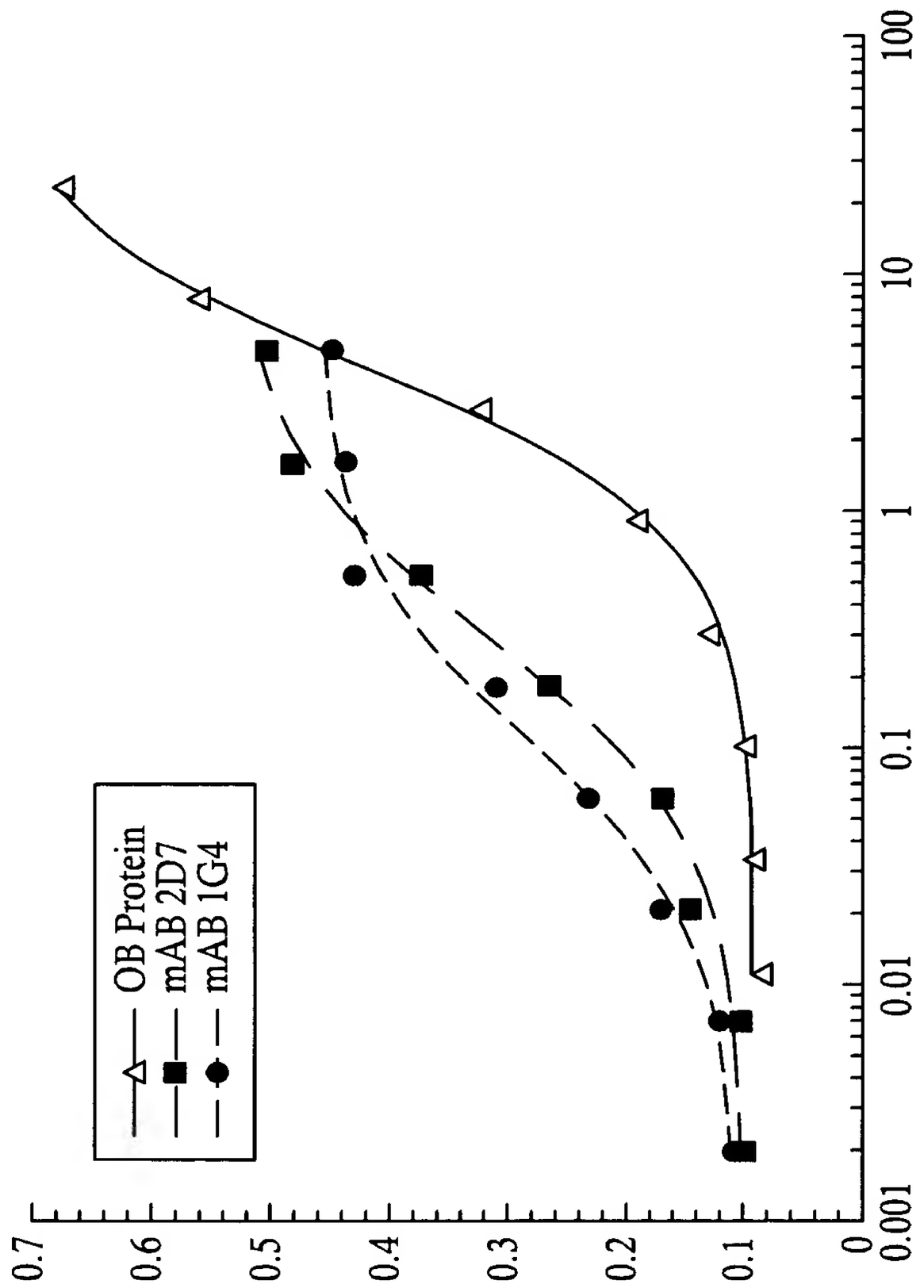

FIG. 18

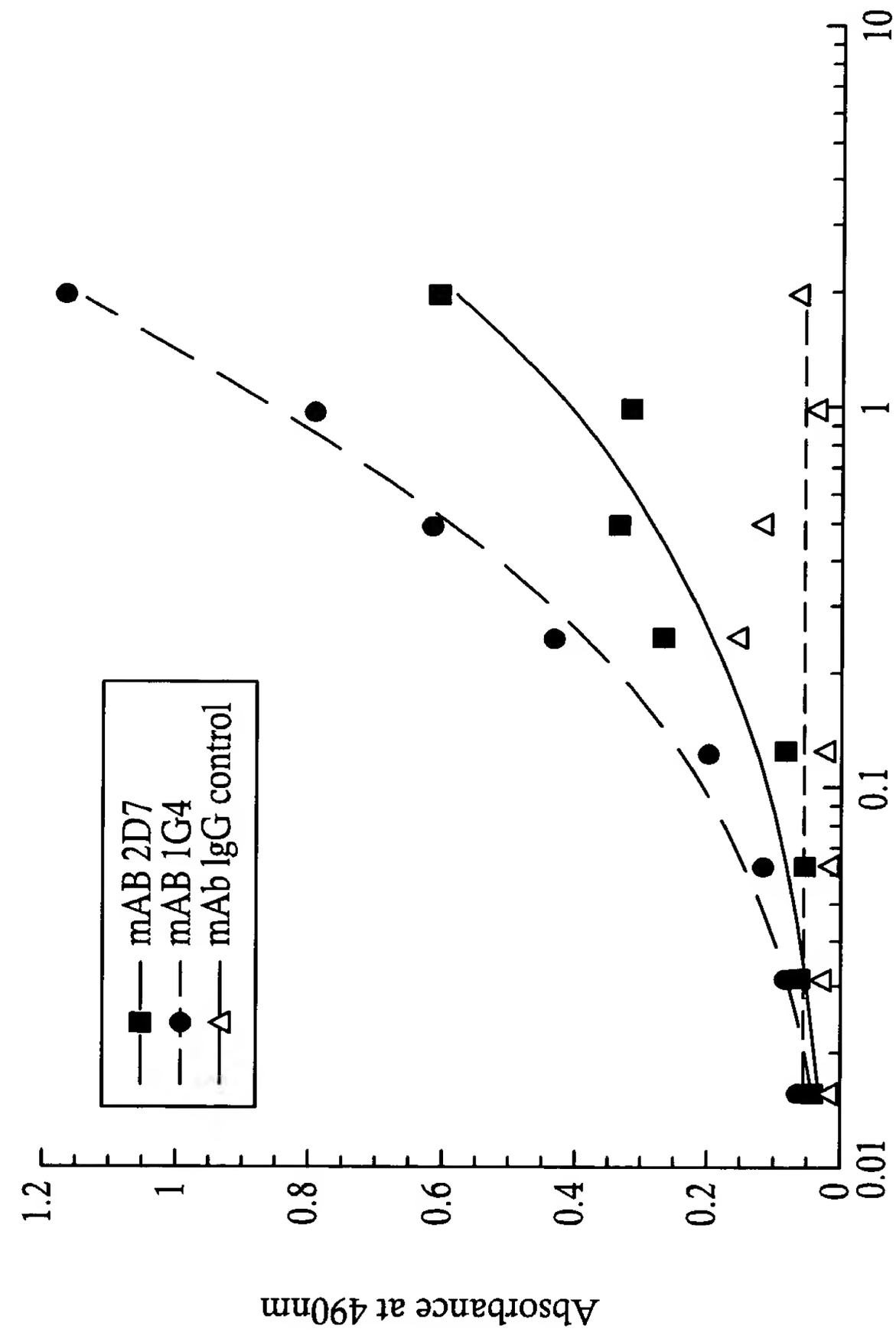

FIG. 19

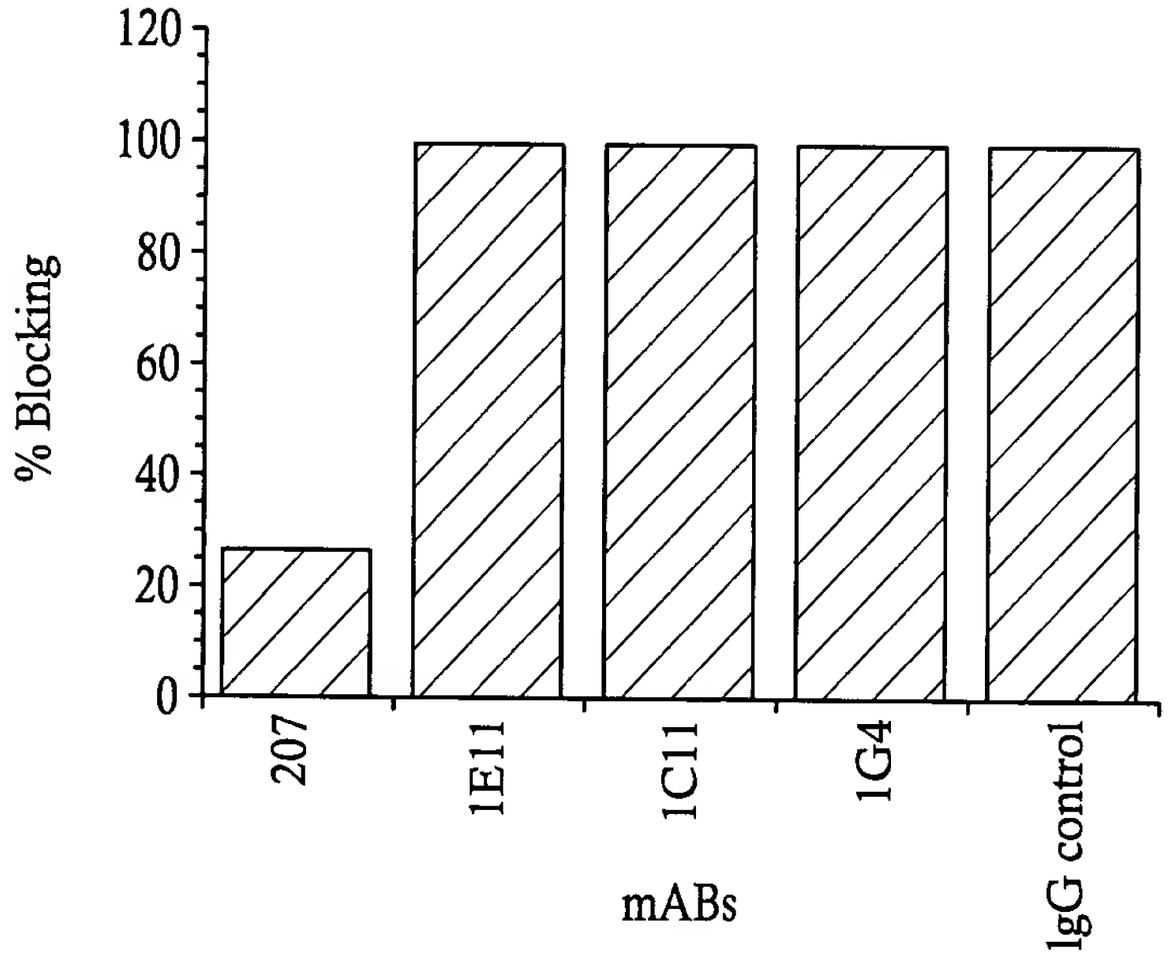

FIG. 20A

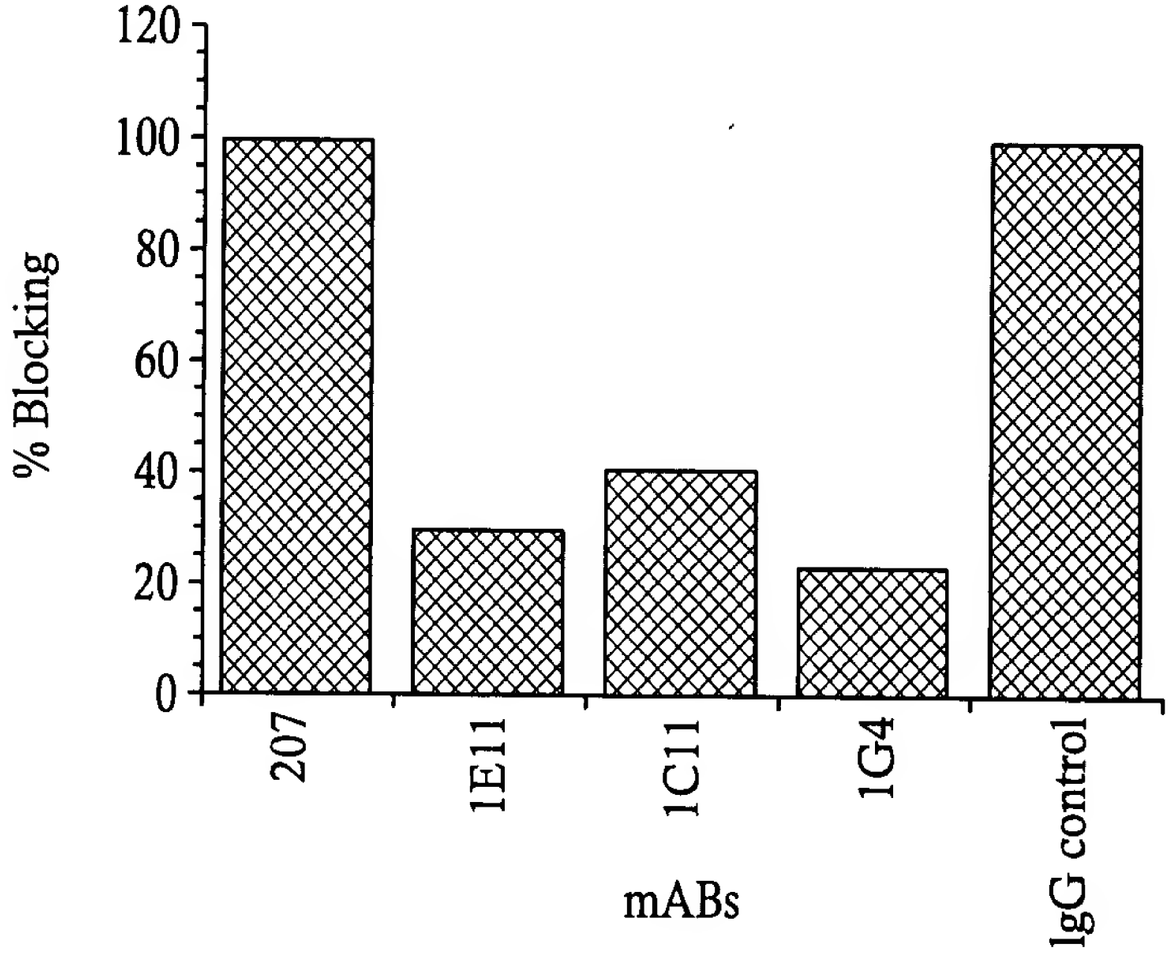

FIG. 20B

```
hWSXR     1  MICQKFCVVLLHWEFIYVITAFNLSYPITPWRFKLSCMPPNSTYDYFLLP
mWSXR     1  MMCQKFYVVLLHWEFLYVIAALNLAYPISPWKFKLFCGPPNTTDDSFLSP

         51  AGLSKNTSNSNGHYETAVEPKFNSSGTHFSNLSKTTFHCCFRSEQDRNCS
         51  AGAPNNASALKGASEAIVEAKFNSSGIYVPELSKTVFHCCFGNEQGQNCS

        101  LCADNIEGKTFVSTVNSLVFQQIDANWNIQCWLKGDLKLFICYVESLFKN
        101  ALTDNTEGKTLASVVKASVFRQLGVNWDIECWMKGDLTLFICHMEPLPKN

        151  LFRNYNYKVHLLYVLPEVLEDSPLVPQKGSFQMVHCNCSVHECCECLVPV
        151  PFKNYDSKVHLLYDLPEVIDDSPLPPLKDSFQTVQCNCSLRG-CECHVPV

        201  PTAKLNDTLLMCLKITSGGVIFQSPLMSVQPINMVKPDPPLGLHMEITDD
        200  PRAKLNYALLMYLEITSAGVSFQSPLMSLQPMLVVKPDPPLGLHMEVTDD

        251  GNLKISWSSSPPLVPFPLQYQVKYSENSTTVIREADKIVSATSLLVDSILP
        250  GNLKISWDSSQTMAPFPLQYQVKYLENS-TIVREAAEIVSATSLLVDSVLP

        301  GSSYEVQVRGKRLDGPGIWSDWSTPRVFTTQDVIYFPPKILTSVGSNVSF
        299  GSSYEVQVRSKRLDGSGVWSDWSSPQVFTTQDVVYFPPKILTSVGSNASF

        351  HCIYKKENKIVPSKEIVWWMNLAEKIPQSQYDVVSDHVSKVTFFNLNETK
        349  HCIYKNENQIISSKQIVWWRNLAEKIPEIQYSIVSDRVSKVTFSNLKATR

        401  PRGKFTYDAVYCCNEHECHHRYAELYVIDVNINISCETDGYLTKMTCRWS
        399  PRGKFTYDAVYCCNEQACHHRYAELYVIDVNINISCETDGYLTKMTCRWS

        451  TSTIQSLAESTLQLRYHRSSLYCSDIPSIHPISEPKDCYLQSDGFYECIF
        449  PSTIQSLVGSTVQLRYHRRSLYCPDSPSIHPTSEPKNCVLQRDGFYECVF

        501  QPIFLLSGYTMWIRINHSLGSLDSPPTCVLPDSVVKPLPPSSVKAEITIN
        499  QPIFLLSGYTMWIRINHSLGSLDSPPTCVLPDSVVKPLPPSNVKAEITVN

        551  IGLLKISWEKPVFPENNLQFQIRYGLSGKEVQWKMYEVYDAKSKSVSLPV
        549  TGLLKVSWEKPVFPENNLQFQIRYGLSGKEIQWKTHEVFDAKSKSASLLV

        601  PDLCAVYAVQVRCKRLDGLGYWSNWSNPAYTVVMDIKVPMRGPEFWRIIN
        599  SDLCAVYVVQVRCRRLDGLGYWSNWSSPAYTLVMDVKVPMRGPEFWRKMD

        651  GDTMKKEKNVTLLWKPLMKNDSLCSVQRYVINHHTSCNGTWSEDVGNHTK
        649  GDVTKKERNVTLLWKPLTKNDSLCSVRRYVVKHRTAHNGTWSEDVGNRTN

        701  FTFLWTEQAHTVTVLAINSIGASVANFNLTFSWPMSKVNIVQSLSAYPLN
        699  LTFLWTEPAHTVTVLAVNSLGASLVNFNLTFSWPMSKVSAVESLSAYPLS

        751  SSCVIVSWILSPSDYKLMYFIIEWKNLNEDGEIKWLRISSSVKKYYIHDH
        749  SSCVILSWTLSPDDYSLLYLVIEWKILNEDDGMKWLRIPSNVKKFYIHDN

        801  FIPIEKYQFSLYPIFMEGVGKPKIINSFTQDDIEKHQSDAGLYVIVPVII
        799  FIPIEKYQFSLYPVFMEGVGKPKIINGFTKDAIDKQQNDAGLYVIVPIII

        851  SSSILLLGTLLISHQRMKKLFWEDVPNPKNCSWAQGLNFQKRTDILL
        849  SSCVLLLGTLLISHQRMKKLFWDDVPNPKNCSWAQGLNFQKRTDTLL
```

FIG. 21

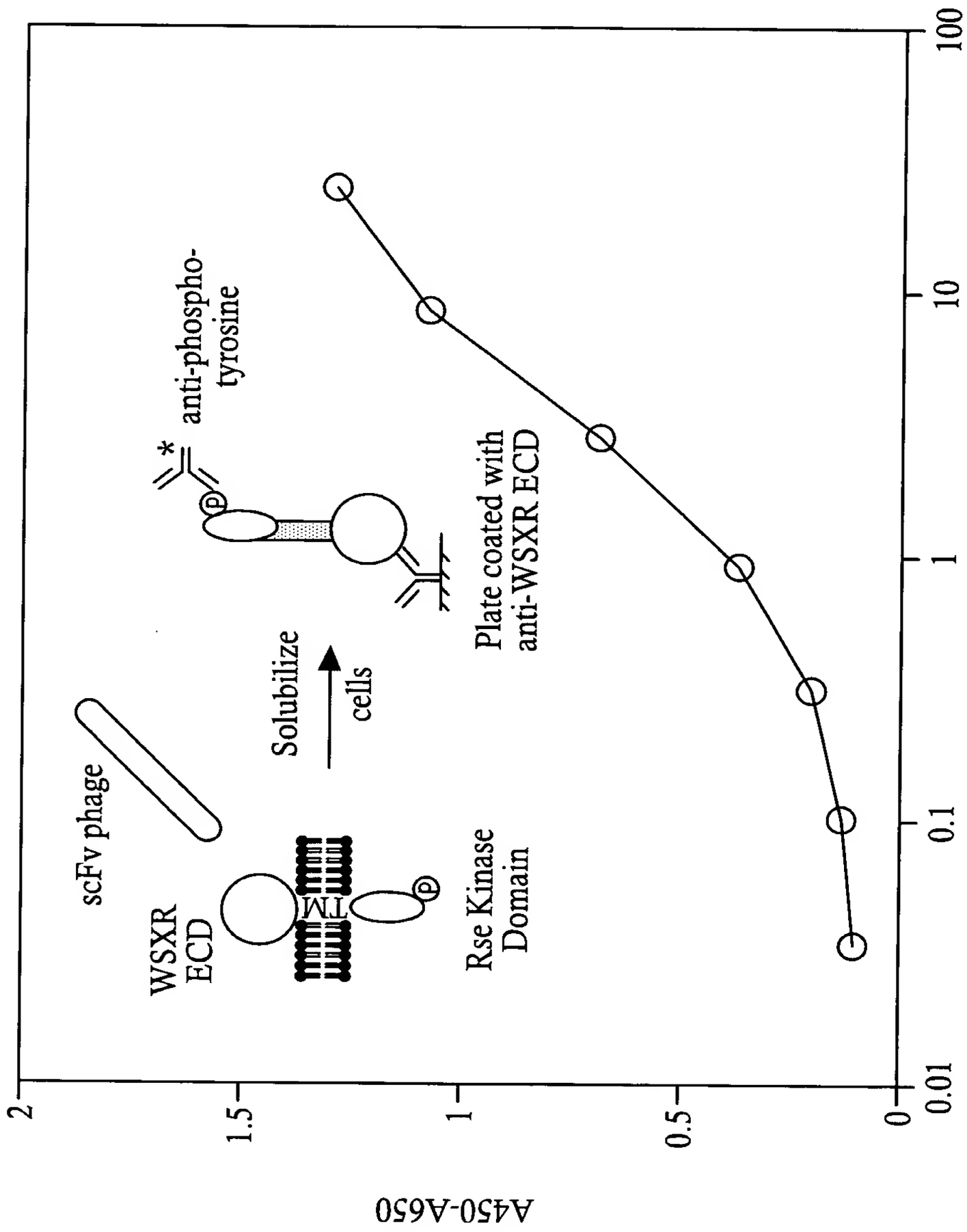

FIG. 22

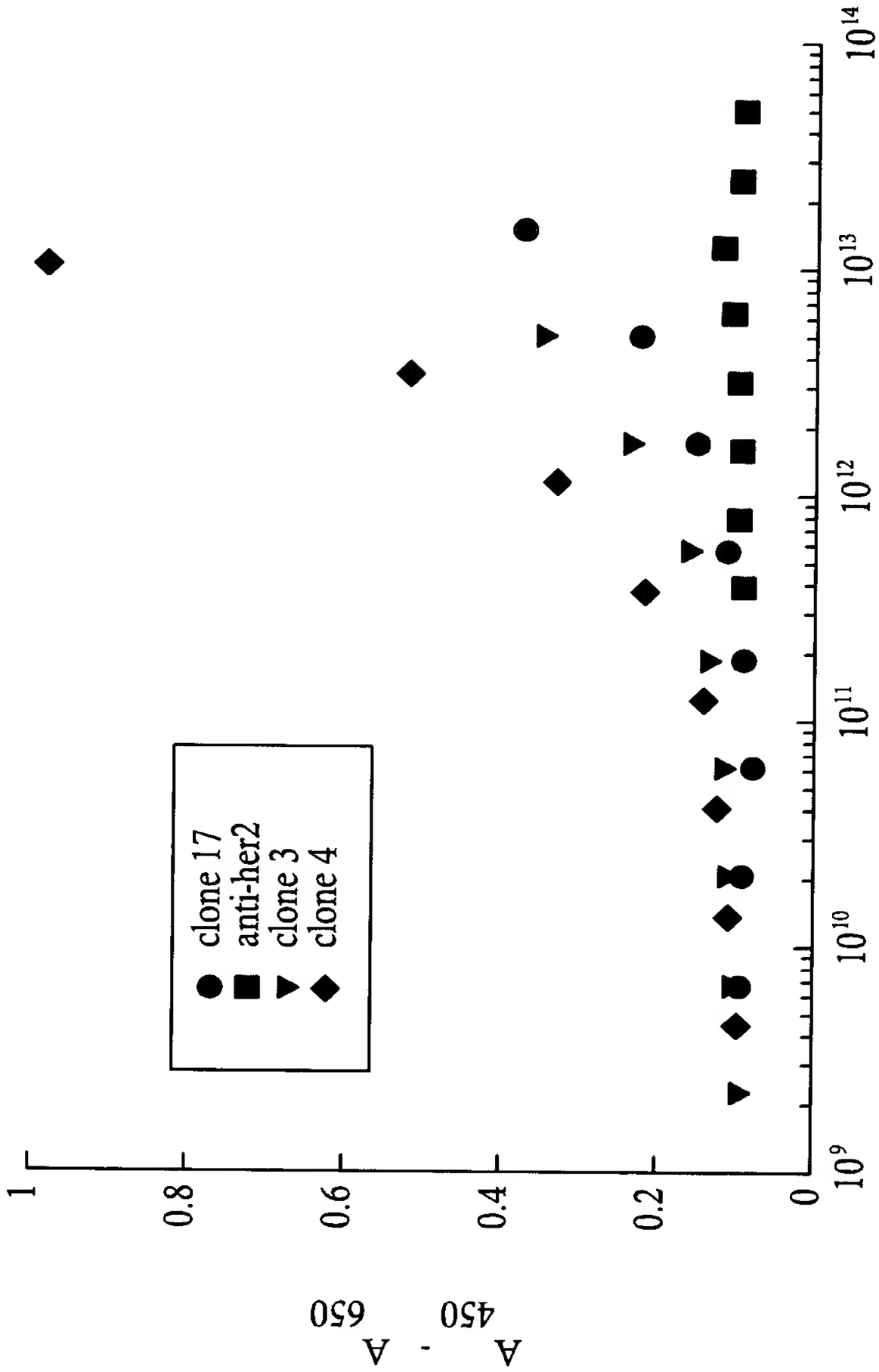

FIG. 23

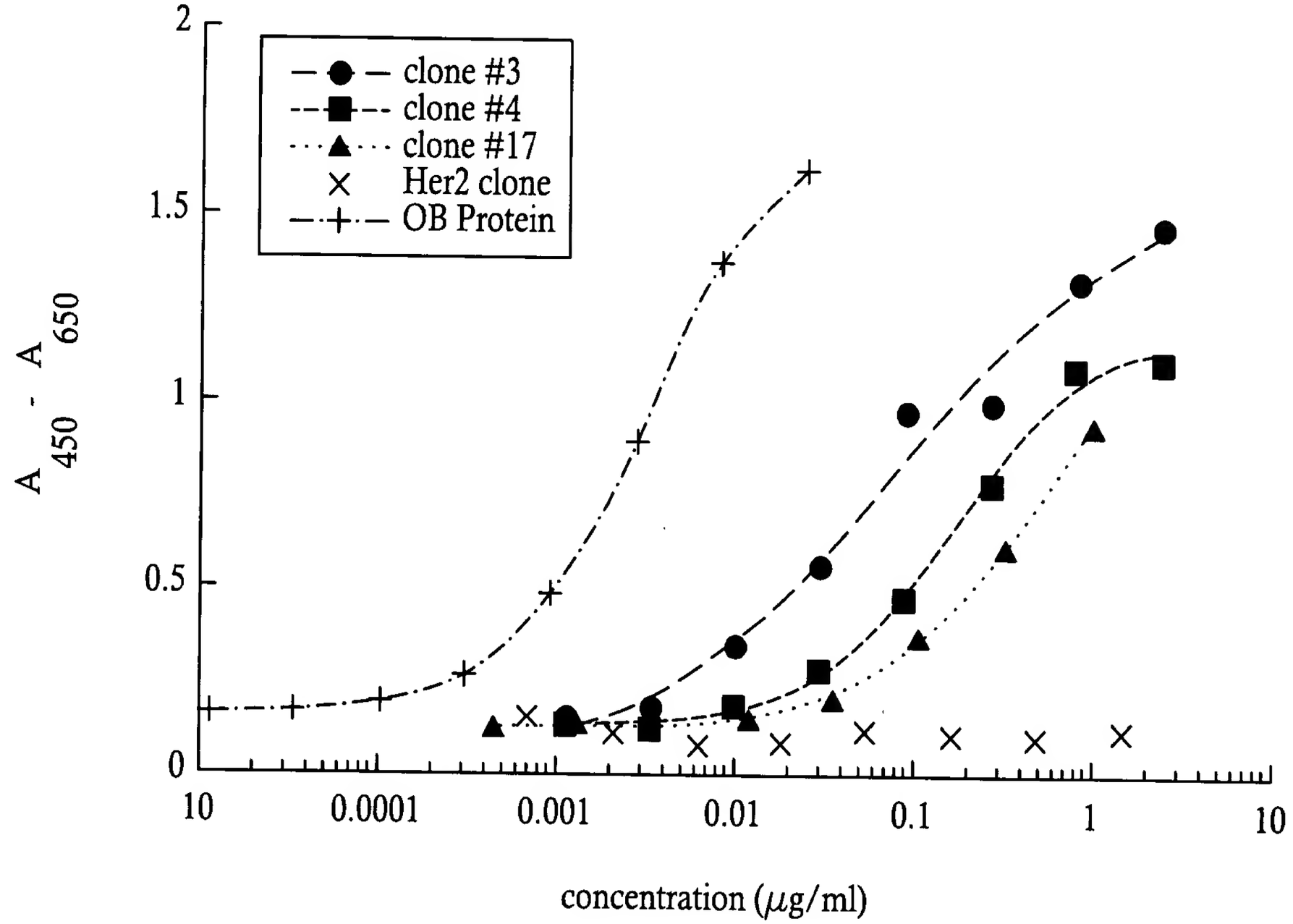

FIG. 24

```
17.scfv      1  QVRLQQSGGGLVQPGRSLRLSCAASGRFTDDYAMHWVRQAPGKGLEWVSG
3.scfv       1  EVQLVQSGAEVKKPGASVKVSCKASGYTFTGYYMYWVRQAPGQGLEWMGW
4.scfv       1  EVQLVQSGAEVKKPGESLKISCQGSGFTFSSYKMNWVRQAPGKGLEWMGG

                                              CDR H1

17.scfv     51  MTWNSGSIGYADSVKGRFTISRDNAKNSLYLQMNSLRAEDTAVYYCAREP
3.scfv      51  INPNSGGTNYAQKFQGRVTMTRDTSIGTAYMELSRLSSDDTAVYYCARDR
4.scfv      51  IIPIFGTANYAQKFQGRVTITADESTSTAYMELSSLRSEDTAVYYCARDR

                    CDR H2

17.scfv    101  HNTDA--------FDIWGRGTLVTVSSGGGGPGGGGSGGGGSDVVMTQSP
3.scfv     101  YYGSSAYHRGSYYMDVWGRGTLVTVSSGGGGTGGGGSGGGGS-SELTQDP
4.scfv     101  VVVPATSLRGG--MDVWGQGTTVTVSSGGGGSGGGGSGGGGSQSVLTQPA

                  CDR H3

17.scfv    143  SFLSAFVGDTITITCRASQ---GIYNYLAWYQQKPGKAPKLLIYAASTLQ
3.scfv     150  A-VSVALGQTVRITCQGDS--LRSY-YASWYQQKPGQAPVLVIYGKNNRP
4.scfv     149  S-VSGSPGQSITISCTGTSSDVGGYNYVSWYQQHPGKAPKLMIYEGSKRP

                                    CDR L1                CDR L2

17.scfv    190  SGVPSRFSGSGSGTEFTLTISSLQPEDFGTYYCQQLI--SYPLTFGGGTK
3.scfv     196  SGIPDRFSGSSSGNTASLTITGAQAEDEADYYCNSRDSSGNHVVFGGGTK
4.scfv     198  SGVSNRFSGSKSGSTASLTISGLQAEDEADYYCSSYTTRSTR-VFGGGTK

                                                  CDR L3

17.scfv    238  VEIK
3.scfv     246  LTVL
4.scfv     247  LTVL
```

FIG. 25